Plate-forme dédiée aux bases Beilstein, Gmelin et Patent Chemistry Database

Gilles Niel (gilles.niel@enscm.fr) – Janvier 2012

AVERTISSEMENTS

Ce manuel décrit les principales caractéristiques de la plate-forme Reaxys (http://www.**reaxys**.com/*)* qui donne accès à l'interrogation des bases Beilstein, Gmelin et de la Patent Chemistry Database pour des recherches ciblées vers la chimie organique, organométallique, et minérale.

Les modalités d'accès à cette plate-forme sont disponibles sur le site de l'INIST (http://titanesciences.inist.fr/) pour les personnes travaillant dans des unités CNRS de l'Institut de Chimie.

Afin de faciliter son utilisation, ce manuel comporte des renvois, par ex., sous la forme A-II-3, au chapitre et/ou paragraphe auquel l'utilisateur peut se référer pour une meilleure compréhension. Lors d'une utilisation avec un ordinateur, ces renvois fonctionnent par hyperlien. Il n'est pas exempt d'erreurs, fautes et coquilles diverses. La bienveillance du lecteur est sollicitée, ses remarques et suggestions seront les bienvenues.

Enfin je tiens à remercier chaleureusement Jacques Coste, professeur retraité de l'Ecole Nationale Supérieure de Chimie de Montpellier, pour y avoir initié avec passion l'enseignement sur les bases de données en chimie et par conséquent la rédaction de manuels tels que celui-ci.

TABLE DES MATIÈRES

A – INTRODUCTION 3
A-I. Contenu global 3
A-II. Structuration de la base 4
A-III- Les principaux types d'interrogation 6
A-IV- Connexion au serveur/Deconnexion 8
A-V- Présentation générale de l'interface 9
B – INTERROGATION PAR PROPRIÉTÉS 12
B-I- A partir des formulaires prédéfinis (Form-based) 12
B-II- En mode expert (Advanced) 13
C – DESSIN DES STRUCTURES ET SOUS-STRUCTURES 20
C-I- Généralités 20
C-II- Dessiner avec MarvinSketch 21
C-III- Atomes 23
C-IV- Liaisons 24
C-V- Groupements Fonctionnels 27
C-VI- Stéréochimie 27
C-VII- Topologie 28
C-VIII- Groupements variables 31
C-IX- Templates 35
D – INTERROGATION PAR STRUCTURE 37
D-I- Recherche par structure exacte 37
D-II- Recherche par sous-structure 38
D-III- Tri des réponses 40
E – INTERROGATION PAR SCHÉMA RÉACTIONNEL 42
E-I- Dessin des structures 42
E-II- Schéma réactionnel 42
E-III- Spécifications des requêtes par schéma réactionnel 44
E-IV- Réactions sélectives 46
E-V- Réactions multi-étapes 47
E-VI- Tri des réponses 47
E-VII- Plans de synthèse – Aide à la rétrosynthèse 48
F – REQUÊTES CROISÉES 52
G – TRAITEMENT DES RÉSULTATS 54
G-I- Navigation dans les résultats 54
G-II- Croisements de listes de réponses 54
G-III- Sauver/Retrouver une requête, une réponse 54
G-IV- Veille scientifique 54
G-V- Export des résultats 55
ANNEXES 57
Annexe 1 - Basic Indexes 57
Annexe 2 - Beilstein Data Structure 58
INDEX 63

A – INTRODUCTION

A-I. Contenu global

La plate-forme Reaxys® permet l'interrogation simultanée de 3 bases distinctes mais interrogeables via la même interface web :

1. CrossFire Beilstein (composés organiques)
2. CrossFire Gmelin (composés minéraux et organométalliques)
3. Patent Chemistry Database (brevets)

Elle contient environ 4,5 M de citations, 20 M de substances et 31 M de réactions.[1] Le présent manuel a pour but de faciliter l'apprentissage de l'interrogation de cette plate-forme. Une documentation détaillée est accessible par le **Help** du navigateur (Figure 4).

A-I-1- Beilstein

Beilstein (http://www.reaxys.com/info/about-overview) est, historiquement, la première organisation d'information en chimie organique. Depuis 1881, le Beilstein Handbuch (Beilstein papier) rassemble les informations concernant la plupart des composés organiques. La série papier couvre les années 1881-1979 ; la dernière édition est parue en 1984.

La version électronique contient des informations sur les composés organiques et leurs propriétés.

Elle est constituée à partir de trois sources différentes.

* Le Beilstein Handbuch, de la série principale jusqu'au supplément IV couvrant la littérature de 1779 à 1959. 1,1 M de composés et leurs propriétés sont indexées.

* Entre 1960 et 1979, de données de la littérature primaire qui avaient été recueillies pour la constitution du Supplément V: 3 M de composés supplémentaires.

* Depuis 1980, les données sont indexées à partir de la littérature primaire.

Toutes ces informations ont été mises sous forme de données informatisées, les données postérieures, indexées à partir de la littérature primaire, ont été rajoutées et cette indexation des résultats de la littérature se poursuit.

Beilstein a d'abord été accessible **on line** par STN (1988) et Dialog (1989). Depuis 1995 le système CrossFire a permis l'interrogation **in house** de la base Beilstein jusqu'à la récente parution de Reaxys® qui donne accès à la base *via* son navigateur usuel.

La base contient des informations sur les composés organiques, leurs propriétés (jusqu'à 350 propriétés différentes) et les réactions permettant leur préparation. En conséquence, si Beilstein est d'abord une base de molécules et de leurs propriétés physiques et physico-chimiques, elle est aussi une base de données de réactions.

Depuis 1980, les titres, résumés et mots-clés sont indexés ce qui fait de Beilstein CrossFire une base bibliographique, sinon exhaustive, du moins utile. Enfin on peut accéder aux articles originaux sous réserve que les abonnements soient localement à jour. Les brevets ne sont plus enregistrés depuis 1980 dans la base Beilstein mais dans la Patent Chemistry Database (cf A-I-3).

A-I-2- Gmelin

Elle est bâtie sur le même principe que Beilstein mais est orientée vers les produits minéraux et organométalliques. Les premières citations indexées dans le *Gmelin HandBook* remontent 1772; la base couvre plus de 2 M de substances pour le développement de matériaux (nano-matériaux, zéolithes et laser-matériaux, les composés de coordination, verres et céramiques, alliages, matériaux dopés, catalyseurs et minéraux.

Plus de 800 types de propriétés chimiques et physiques sont interrogeables, en particulier les propriétés électriques, magnétiques, thermiques, cristallines et physiologiques.

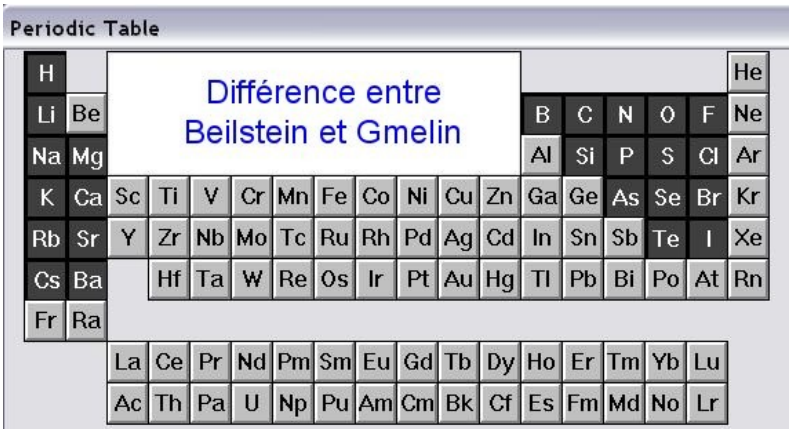
---

1 - On peut actualiser ces chiffres à cette adresse :
https://www-reaxys-com.chimie.gate.inist.fr/documentation/about_query.htm

Jusqu'en 1980 tous les composés organiques et organométalliques ont été indexés dans Beilstein, depuis 1980 de nombreux composés organométalliques sont indexés dans Gmelin ainsi que les organométalliques qui comportent un des atomes d'une des cases non grisées de la Figure 1.

La version électronique de la base n'est pas équivalente à la version papier. La plupart des données numériques, contenues dans le *Handbook* imprimé, jusqu'en 1975 sont intégrées dans la plate-forme Reaxys. Une partie des données, extraites de 110 journaux entre 1976 et 1997 mais non du *HandBook*, est intégrée à Reaxys. De 1995 – date où l'institut Gmelin a fermé – à 2000 – où l'indexation a été reprise par la société MDL – il y a des lacunes dans la couverture des données. Actuellement la base Gmelin est compilée par la Société Allemande de Chimie (GDCh) qui couvre la littérature issue de 62 journaux.

Figure 1 : indexation des substances en fonction des éléments du tableau périodique


Periodic Table
Différence entre
Beilstein et Gmelin
H
He
Li Be
B C N O F Ne
Na Mg
Al Si P S Cl Ar
K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
Rb Sr Y Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I Xe
Cs Ba Hf Ta W Re Os Ir Pt Au Hg Tl Pb Bi Po At Rn
Fr Ra
La Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu
Ac Th Pa U Np Pu Am Cm Bk Cf Es Fm Md No Lr

A-I-3- Patent Chemistry Database

La PCD indexe des substances et des réactions contenues dans des brevets de puis 1976; elle n'est donc pas une base de brevets similaire aux bases de brevets classiques (USPTO, Espacenet, ….) mais fournit un grand nombre de propriétés utiles en synthèse (schémas réactionnels, données spectrales), en biologie (données pharmacologiques et liées aux Relations-Structure-Activité). Les premiers brevets cités apparaissent en 1976 ce qui fait de la PCD une base complémentaire de Beilstein qui cite des brevets sur la période : 1886 - 1980.

En ce qui concerne les substances, la PCD fournit des informations non seulement sur les substances complètement décrites mais aussi sur les substances dites "prophétiques" (traduction littérale), c'est-à-dire celles qui sont revendiquées par les inventeurs comme pouvant être préparées de manière analogue aux substances mais sans préciser les données associées (rendement, propriétés physico-chimiques, …). Ces substances sont souvent représentées par des formules génériques, dites de Markush.

La PCD représente environ 845 000 brevets pour 6,9M de substances et 5,7M de réactions.

A-II. Structuration de la base

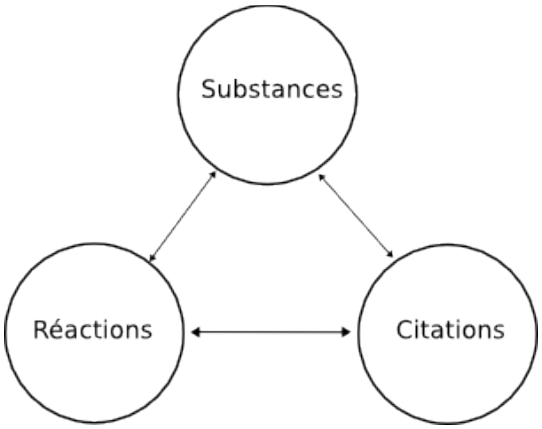
Les données de Reaxys sont indexées dans trois parties distinctes : substances, réactions et citations qui sont reliées par des hyperliens permettant de naviguer entre les 3 parties.

A-II-1- Substances

La partie Substances contient environ 20 M de composés. C'est beaucoup, mais moins que les 64 M de substances indexées par le CAS (Chemical Abstracts Service). Cela est dû à des raisons éditoriales, historiques et économiques. Si jusqu'en 1979, l'exhaustivité est recherchée, ce n'est plus le cas pendant les années 1980.

On peut les rechercher par interrogation structurale exacte ou sous-structurale. Aux substances sont associées leurs propriétés physiques, physico-chimiques, pharmacologiques, etc. Jusqu'à 350 propriétés différentes sont indexées pour Beilstein et 800 pour Gmelin. Ces propriétés sont indexées avec des valeurs soit numériques soit textuelles.

Les composés indexés doivent répondre à certains critères; ils doivent être purs, suffisamment bien décrits et de structure prouvée.

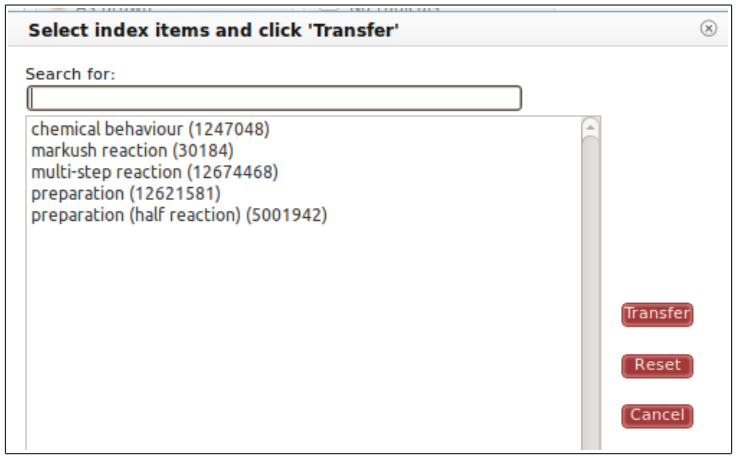


Les peptides, jusqu'à 12 résidus d'acides aminés, et les polysaccharides (6 résidus) sont indexés avec leur structure. Les peptides et polysaccharides de plus grande taille et les autres biomolécules sont seulement enregistrés par leur nom en tant que biomolécules.

Les polymères, quoique peu nombreux, sont également seulement indexés par leur nom avec, le cas échéant le monomère sous forme structurale.

A-II-2- Réactions

La partie Reactions contient environ 31 M de réactions qui se répartissent comme montré sur la copie d'écran ci-dessous (septembre 2011) :



Remarque : La classification ci-contre peut être obtenue en interrogeant par propriété – dans ce cas le champ RX.CL (cf. section B-II-1).

Les réactions de préparation (≈ 12,6 M) ne sont enregistrées que si elles sont adaptées à des synthèses à grande échelle. Les méthodes à faibles rendements ne sont indexées que si elles ont une utilité préparative.

Elles sont interrogeables par structure exacte ou sous-structure et certaines propriétés leur sont associées : le rôle (réactant, produit), la correspondance entre les atomes du(es) réactant(s) et du(es) produit(s), les modifications de liaisons, les centres réactionnels et, le cas échéant, le résultat stéréochimique (rétention, inversion) de la réaction.

La plupart des réactions sont enregistrées en tant que réaction de préparation des composés qui sont indexés dans la base. On trouve également des réactions concernant la réactivité ou **chemical behaviour** (≈ 1,2 M). Jusqu'en 2008, la base contenait principalement des réactions mono-étape : c'est, au moins au départ, la philosophie de constitution du Beilstein *HandBook* qui enregistre les réactions de préparation comme une propriété des composés, c'est à dire seulement la dernière étape de la synthèse.

La politique d'indexation ayant changé en 2008, on trouve aussi de nombreuses réactions multi-étapes (≈ 12,7 M). Elles sont enregistrées et notées **multi-step reaction**; un exemple est détaillé en section E-V. Réaction mono- ou multi-étapes, l'isolement et la purification des produits intermédiaires sont décrits.

On note enfin l'existence d'un nombre important de demi-réactions (**half reactions)** c'est-à-dire une réactions où seul le produit (ou le réactant) est connu (≈ 5 M).

NB : Par conséquent, une interrogation par schéma réactionnel complet :

Réactant ⟶ Produit

peut ne pas aboutir si la réaction n'est décrite que sous forme de **half reaction**. De plus, dans les indexations de ces réactions, les centres réactionnels, liaisons cassées, etc, ne sont pas indiqués : ces spécifications ne devront donc pas être mentionnées dans les requêtes.

A-II-3- Citations

La partie Citations contient environ 4,5 M de citations pour ≈ 35 M rapportées dans le CAS. L'explication tient au nombre de journaux indexés, environ 1500 pour la base CAPlus du CAS, 400 pour Reaxys. Si le nombre de journaux actuellement analysés pour constituer la base est relativement important, cela n'a pas toujours été le cas dans les années antérieures. En effet, jusqu'en 1980, la couverture de la littérature était très importante; le nombre de journaux est passé de 85 dans les années 1980 à 120 après 1990 pour atteindre les 400 depuis quelques années. Certains journaux ont été totalement absents pendant 10 ans.

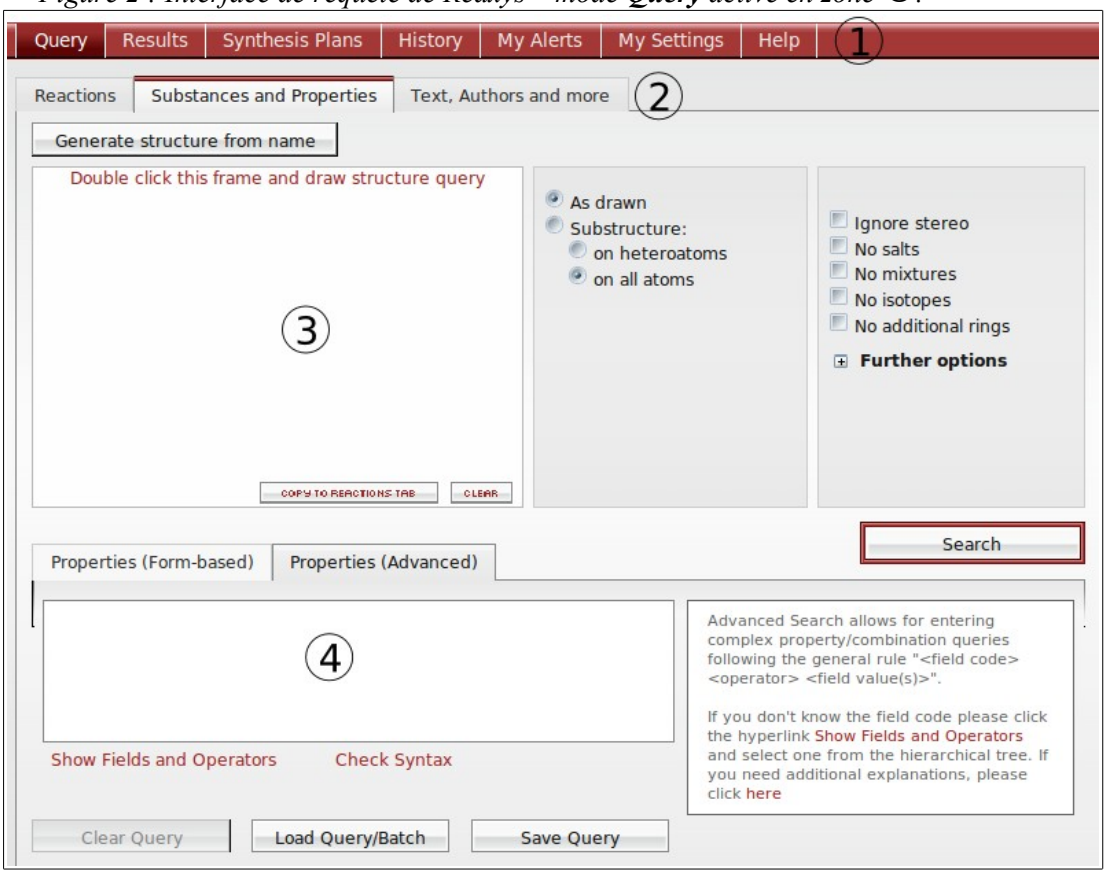
Depuis 1980, les résumés, titres et mots-clés sont également indexés.

A-III- Les principaux types d'interrogation

L'interface actuelle propose plusieurs modes de requête regroupés sur la même fenêtre (Fig. 2) :

- par structure (onglet **Substances and Properties**)
- par schéma réactionnel (onglet ouvert par défaut) } zone ② de la Fig. 2
- de type bibliographique (onglet **Text, Authors and more**)
- par propriété (onglets **Properties**) zone ④ de la Fig. 2
- par croisement des modes précédents (cf. section F).

*Figure 2 : Interface de requête de Reaxys – mode **Query** activé en zone ①.*



A-III-1- Interrogation par structure

Deux possibilités d'interrogation: par structure exacte et par sous-structure.

Les structures sont dessinées dans un éditeur de molécules (cf. section C) puis transférées dans la zone ③ de la Fig. 2 en mode requête. Pratique détaillée en section D.

A-III-1-a- Interrogation par structure exacte

Pour les substances, l'interrogation par structure exacte est le moyen le plus sûr étant donnés :

- les possibilités d'indexations multiples
- les problèmes de nomenclature

Toutefois l'outil **Generate structure from name** peut être d'une grande utilité car plus rapide.

A-III-1-b- Interrogation par sous-structure

Ce moyen très puissant d'interrogation est un des avantages que présentent les bases de données électroniques par rapport à la documentation papier.

A-III-1-c- Problèmes d'indexation des réactants, réactifs et catalyseurs

Les **reactants**, **products**, **reagents** et **catalysts** ne sont pas indexés de façon univoque (cf B-II-4).

A-III-2- Interrogation par Schéma Réactionnel

Hormis les **half-reactions**, l'ensemble des réactions de la base sont interrogeables par schéma réactionnel complet. Pratique détaillée en section E.

Lors d'une interrogation par schéma réactionnel seuls les produits de départ (**reactants**) et produits de la réaction (**products**) doivent figurer dans l'équation de réaction. Les catalyseurs, réactifs et solvants n'apparaissent pas à ce niveau : ils sont indexés sous forme texte et doivent être spécifiés, le cas échéant, en tant que faits, dans l'un des deux onglets intitulés **Conditions** de la Fig.7 (cf. B-II-1).

L'interrogation par sous-structure est un mode très performant d'utilisation des bases de données de réactions. Ce type d'interrogation implique que, dans la base, les schémas réactionnels soient indexés avec :

- les indications de correspondance des atomes entre réactant et produit (**atom to atom mapping**)
- les centres de réaction
- les modifications de stéréochimie

NB : On constate cependant que environ 30 % des réactions ne sont pas mappées et que pour certaines autres réactions le mapping n'est pas chimiquement pertinent.

A-III-3- Interrogation de type bibliographique

L'ensemble des citations est interrogeable par auteur, journal ou brevet. A partir de 1980, les résumés, titres et mots-clés sont également indexés dans Beilstein.

L'accès aux articles originaux se fait à partir du lien **Full Text** par un résolveur de liens local quel que soit le contexte (citations, substances ou réactions). On remarquera aussi la possibilité de visualiser les références citantes d'un article par le biais du portail Scopus.

A-III-4- Interrogation factuelle

Ces interrogations peuvent être effectuées par valeur numérique ou mot-clé dans l'un ou plusieurs des très nombreux champs de la base Reaxys (≥ 800). Pratique détaillée en section B.

A-III-4-a- Types des données contenues dans la base

- les composés (nom, formule brute, formules partielles, n° CAS)
- les données bibliographiques
- les caractéristiques des réactions (réactifs, catalyseurs, solvants, temps de réaction, température, pression, rendement, etc)
- les propriétés physiques, physico-chimiques,
- les propriétés toxicologiques, pharmacologiques et, plus récemment, écologiques.

A-III-4-b- Informations données dans les réponses

Pour chaque type de données sont fournies les références bibliographiques ainsi que des valeurs numériques pour certains types seulement. En fonction de la période d'indexation, les valeurs numériques

sont ou ne sont pas données.

* Pour la partie de la base qui correspond à des données publiées entre 1779 et 1959 (1,1 M de composés et leurs propriétés), les valeurs numériques des propriétés physiques sont données.

* Entre 1960 et 1979, 3 M de composés supplémentaires sont enregistrés dans Beilstein. Pour les points de fusion, ébullition, densité, indice de réfraction, pouvoir rotatoire, isolement des produits naturels et dérivés chimiques, des valeurs numériques sont indexées. Pour toutes les autres propriétés physiques et chimiques seules les références de la littérature sont indexées. Cette partie est actuellement complétée afin de contenir des données pour tous les champs.

* Depuis 1980, les données sont indexées à partir de la littérature primaire et des informations détaillées sont données pour toutes les propriétés physiques et chimiques. Les données pharmacologiques et écologiques sont également indexées.

A-III-4-c- Problèmes d'indexation

Il faut tenir compte de plusieurs problèmes qui sont développés en section B-II-4

- L'indexation des composés réactifs et catalyseurs n'est pas univoque
- Les définitions des réactants, réactifs et catalyseurs ne sont pas toujours suivies
- La possibilité d'utiliser des **basic indexes** permet de pallier certains de ces problèmes.

A-III-5- Interrogation croisée : Structure ou Schéma Réactionnel et Faits

On peut aussi interroger par structure ou schéma réactionnel prenant en compte des contraintes de réactif (catalyseur, solvant) qui seront précisées en mode **Query** dans la partie **Conditions** (cf. B-II-1). Ce type d'interrogation est possible à propos de tous les faits (ou champs) indexés dans la base (auteurs, t° de réaction, rendements, année de publication, etc).

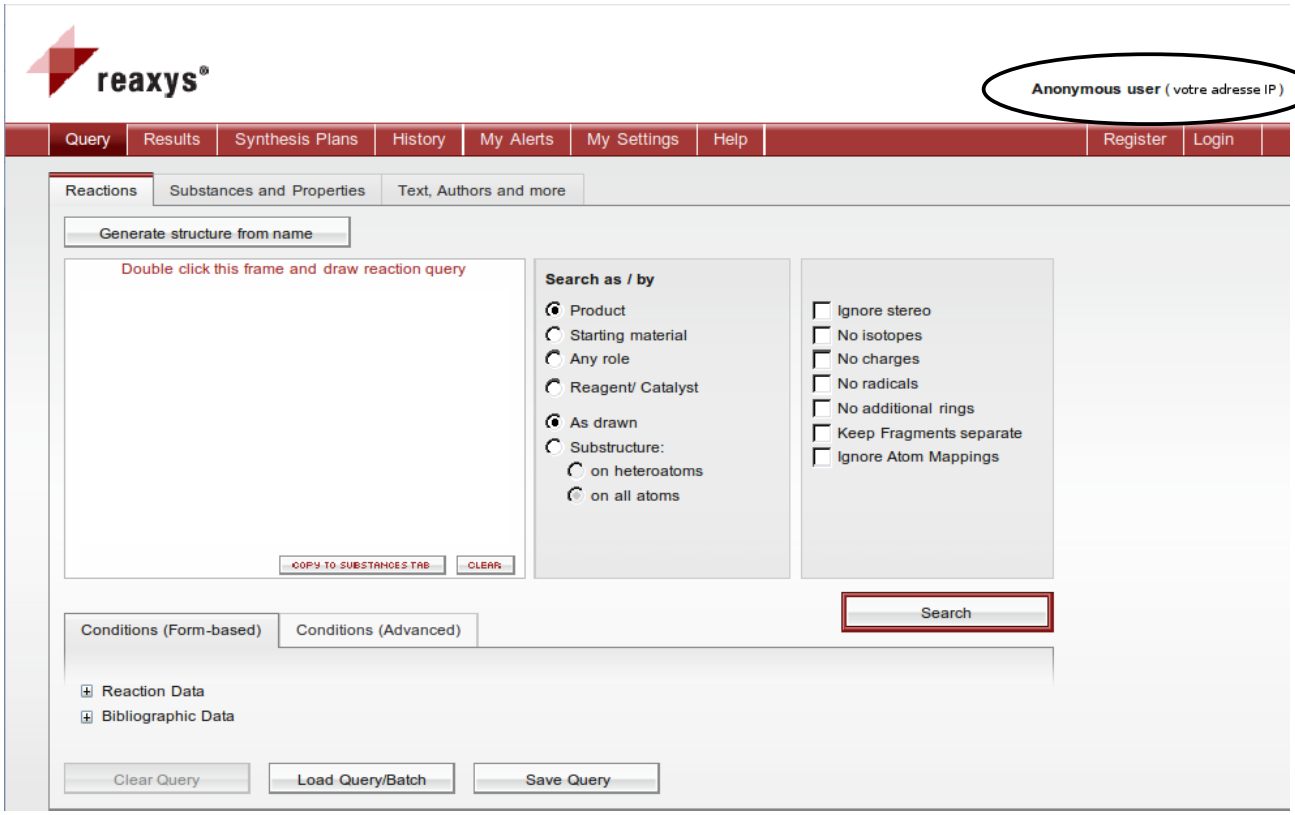
A-IV- Connexion au serveur/Deconnexion

La plate-forme Reaxys fonctionne comme une application Web accessible *via* Internet (https://www-reaxys-com.chimie.gate.inist.fr/reaxys/). Les personnels, travaillant dans des unités reconnues par le département Chimie du CNRS, peuvent accéder à cette plate-forme même s'ils ont un statut universitaire ou s'ils sont non permanents.

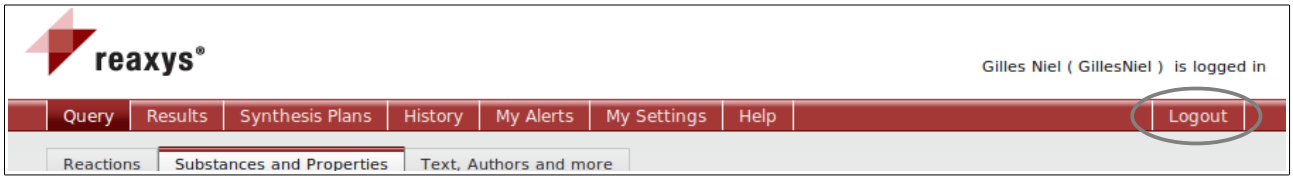
La Figure 3 montre la page d'accueil de la plate-forme, ouverte par défaut sur l'utilisateur **Anonymous user** suivi de votre adresse IP (en haut à droite de la figure).

Il est conseillé de se connecter *via* le bouton **Login** qui permet à chaque utilisateur de s'identifier et de rentrer dans son espace personnel; dans son espace personnel, il est possible de sauver ses préférences (**My Settings**), ses requêtes et listes de réponses (**History**) et de se créer des alertes sur un sujet déjà interrogé (**My Alerts**). Ce type de connexion est possible dans l'enceinte de votre institution (université, …).

Figure 3 : Page d'accueil de la plate-forme Reaxys



Après connexion on observe comme seule modification la présence du bouton **Logout** dans le bandeau principal de la copie d'écran ci-dessous.



N.B. Les requêtes et les listes de réponses sont effacées si elles n'ont pas été sauvegardées pendant la session d'interrogation (cf. section).

A-V- Présentation générale de l'interface

L'interface de Reaxys s'inspire de celle du *Beilstein Commander,* dans sa version V7. Elle est compacte (Fig. 4), intuitive et présente sur une seule page tous les outils qui vont être nécessaires pour poser des requêtes, consulter les réponses, ….

Dans le bandeau principal (zone ①), sont disponibles les différents modes :

- **Query** pour lancer les interrrogations qui se décline, dans la zone ② en 3 onglets (réactions, substances et propriétés, citations)
- **Results** pour afficher et pour traiter les réponses obtenues (cf. section G)
- **Synthesis Plans** qui permet d'afficher des plans de synthèse à partir d'une substance qu'elle soit réactant ou produit d'arrivée (cf. E-VII)
- **History** pour visualiser ET sauvegarder les requêtes ET les listes de réponses (cf. G-II, G-II)
- **My Alerts** outil de veille technologique (cf. G-IV)
- **My Settings** pour modifier ses préférences (choix de l'éditeur de structures, des options de recherche pour les substances et réactions, ….), les données de son profil personnel ou encore changer son mot de passe.
- **Help** : aide interactive au format html.

Cette zone ① ne change pas lorsque l'on change de mode : **Query, Results**, …

Nous avons déjà vu la zone ② dans le paragraphe A-III où l'on choisit le contexte de la requête (réactions, substances et propriétés, citations).

La zone ③ **(Generate structure from name)** permet de retrouver une structure dans la zone ④ en la construisant automatiquement à partir de son nom ou son code InChI, ou son code SMILES ou son numéro CAS.

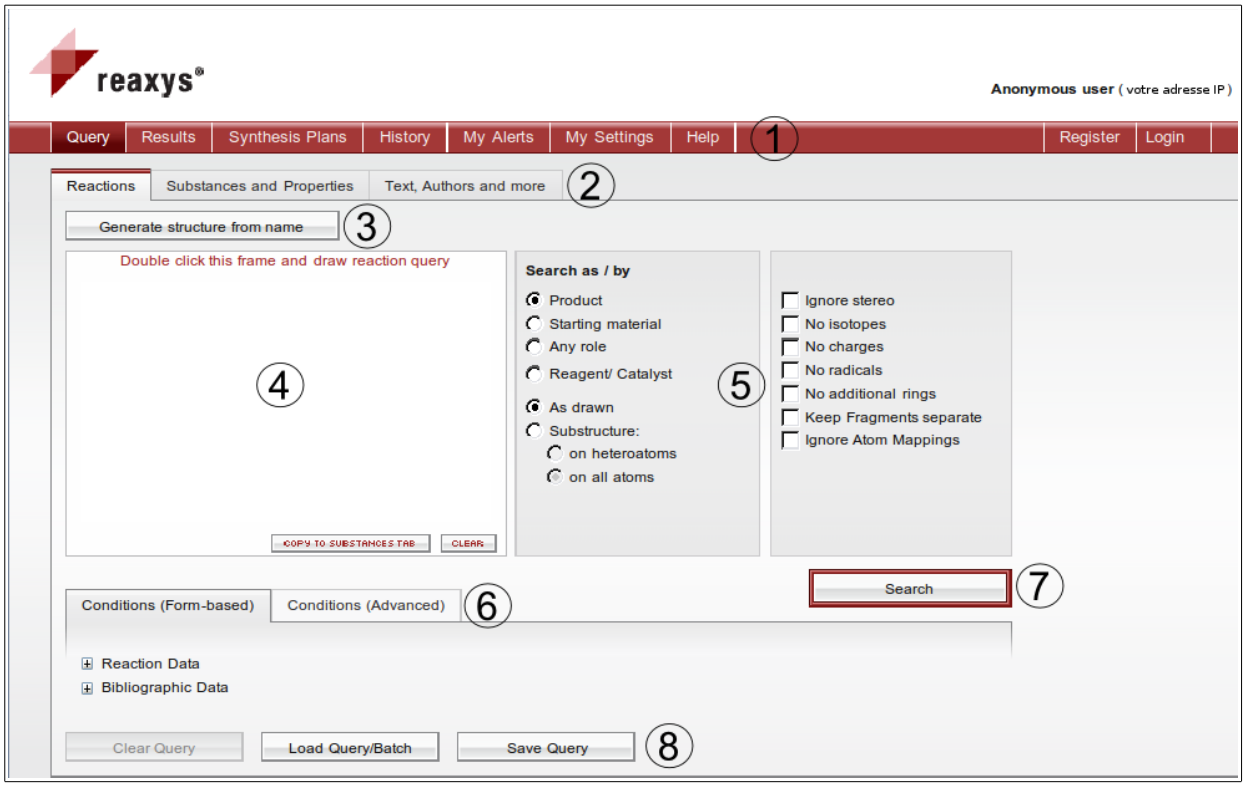
Les options de requête sont personnalisables dans la zone ⑤.

La zone ⑥ varie selon le contexte de la requête mais propose dans chaque cas une recherche guidée par un formulaire de requête ou une recherche en mode expert.

Enfin la zone ⑧ permet de sauver une requête localement (qui apparaîtra donc dans l'historique) ou de charger une requête, précédemment sauvegardée, voire un lot de requêtes.

*Remarque : une force de cet interface en mode **Query** est qu'il permet de réaliser des interrogations simultanément en dessinant une structure ou sous-structure (zone ④) ET en interrogeant par propriétés (zone ⑥).*

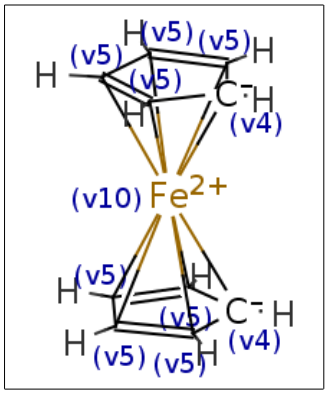
Figure 4 : présentation de l'interface



A-V-1- Domaine de recherche

Les questions sont posées à toute la plate-forme Reaxys, sans distinction de l'origine de la base : Beilstein, Gmelin ou Patent Chemistry Database. C'est l'option par défaut et elle est implicite.

Pour des composés organométalliques qui étaient référencés à la fois dans Beilstein et Gmelin, on observe maintenant une homogénéisation dans la représentation des structures de ces substances. C'est par ex. le cas pour le ferrocène (cf. schéma ci-dessous).



Enfin, on peut restreindre le domaine d'interrogation à une liste de réponses précédemment sauvegardée, ou à une liste de réponses obtenue dans la session courante. Les détails pratiques sont explicités aux paragraphes G-II.

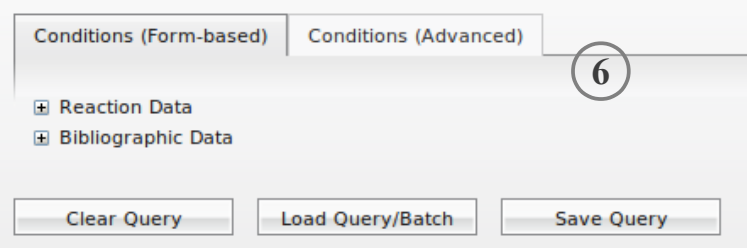
B – INTERROGATION PAR PROPRIÉTÉS

Les interrogations par propriétés sont un moyen rapide d'obtenir une information précise sur une substance, une réaction ou une citation. C'est une des particularités qui sont historiquement un des atouts majeurs de Reaxys par rapport au CAS. Suivant le contexte de la requête (réaction, substance ou citation), il sera possible d'accéder aux propriétés à partir de la zone ⑥ de la Fig. 4 ou 5, si l'on s'intéresse aux réactions.

Remarque : l'ensemble des propriétés, décrites dans l'arborescence complète de la base, est donnée en Annexe 2.

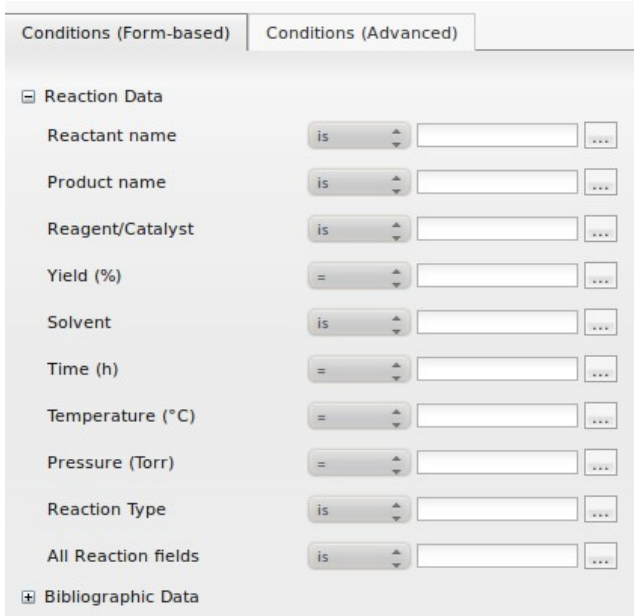
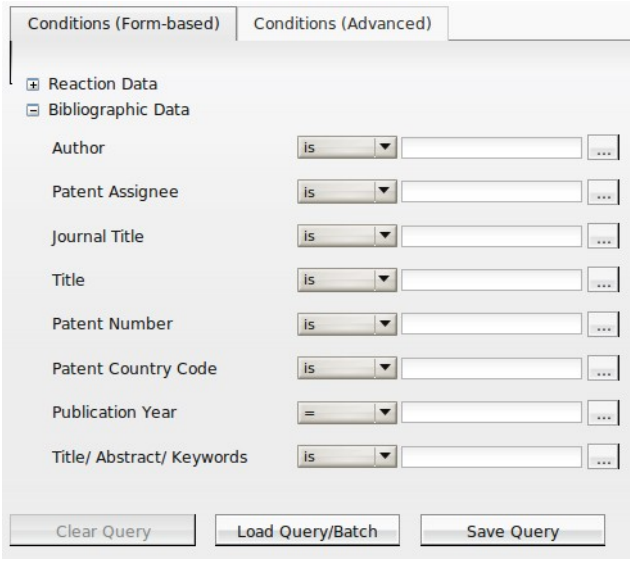
Il faudra choisir l'un des deux onglets **Conditions (Form-based)** ou **Conditions (Advanced)** - la Fig. 5 – qui permettent d'interroger à partir de formulaires prédéfinis (cf. B-I) ou en mode expert (cf. B-II).

Figure 5 : Arborescence des propriétés concernant les réactions



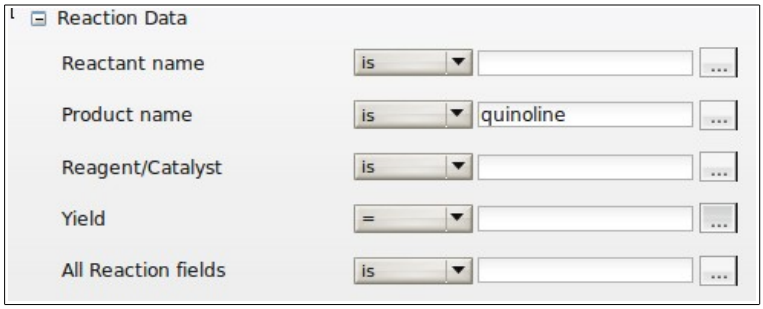
B-I- A partir des formulaires prédéfinis (Form-based)

Si l'on choisit l'onglet **Conditions (Form-based)**, on pourra développer l'item **Reaction Data** ou **Bibliographic Data** dont les copies d'écran sont montrées ci-dessous.

Quelques propriétés des réactions	et les citations associées
	

Ce type d'interrogation est moins précis que si l'on avait dessiné un schéma réactionnel complet (sujet abordé en section E-II-2) mais peut être utile pour connaître par exemple l'ensemble des réactions conduisant à un composé connu. Supposons que l'on cherche les produits de départ qui permettent de synthétiser la quinoléine, on saisira **quinoline** dans le champ correspondant, ici **Product name**.

La requête, montrée ci-dessous, conduit à 399 réactions de préparation.



Si l'on s'intéresse à des substances, on pourra développer l'item **Substance Data** de la Fig. 6a puis l'item **Physical Data** de la Fig. 6b pour aboutir à un choix multiple permettant d'interroger les propriétés les plus usuelles pour une substance.

Figure 6a

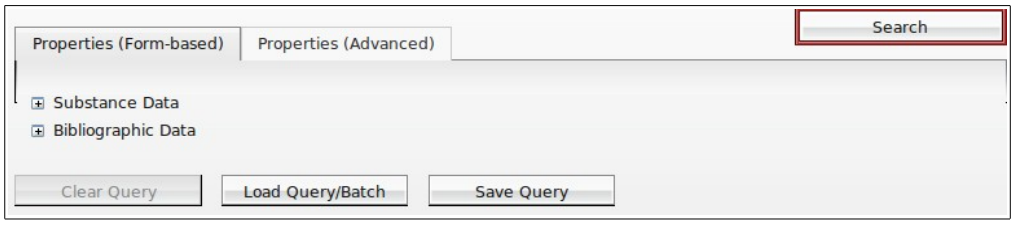
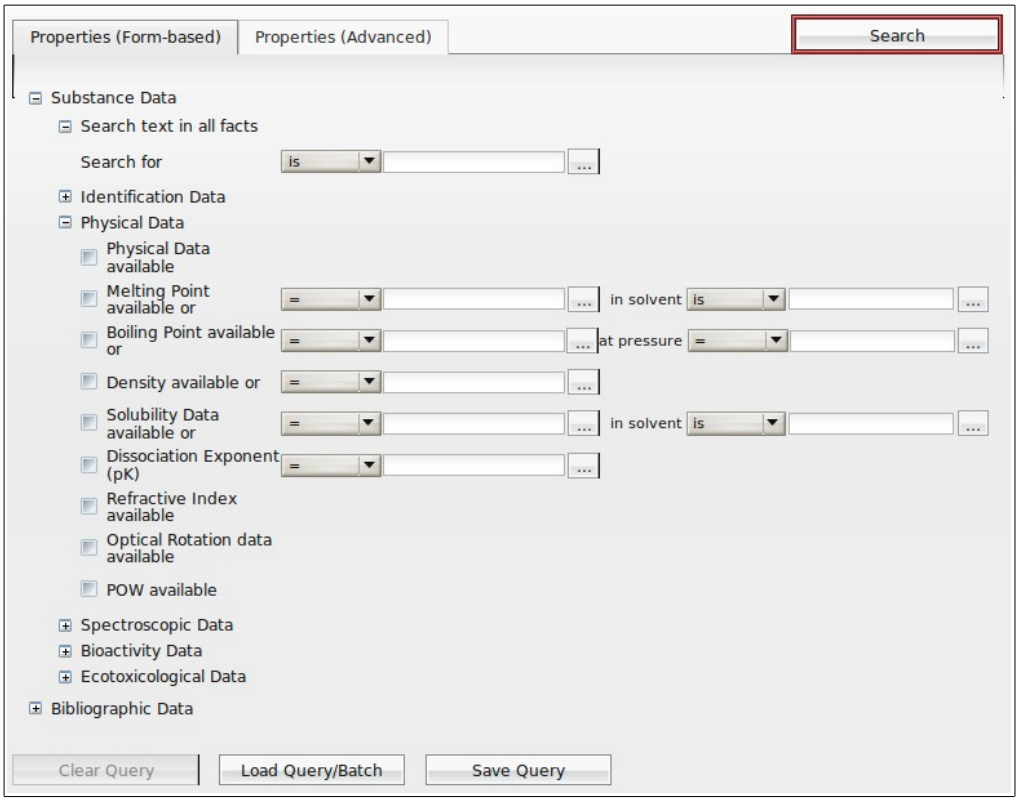


Figure 6b



Pour des requêtes plus compliquées ou pour faire apparaître une propriété absente des formulaires précédents, il faudra utiliser le deuxième onglet : **Properties (Advanced)**.

B-II- En mode expert (Advanced)

B-II-1- Requête générale

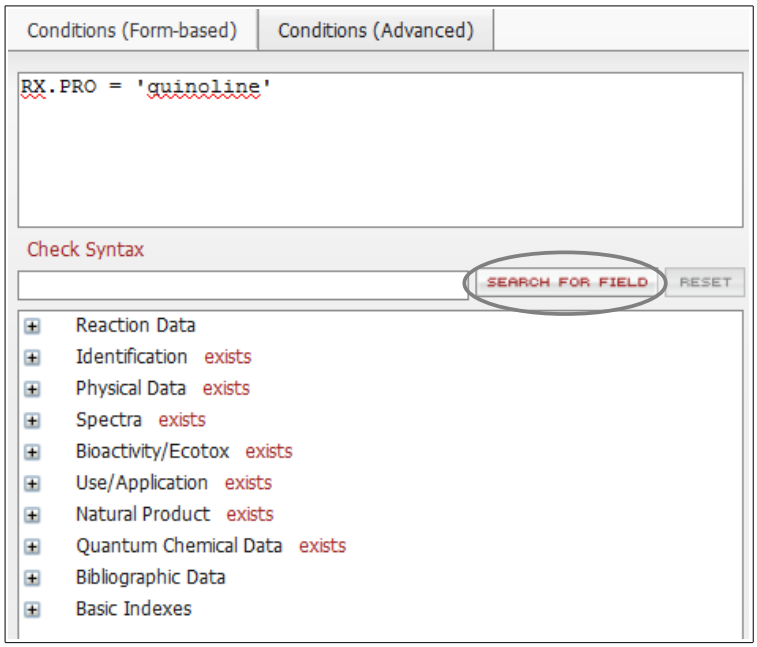
Si l'on choisit l'onglet **Properties (Advanced)**, on accédera à l'ensemble des propriétés de Reaxys à condition de connaître la syntaxe exacte à utiliser, c'est-à-dire :

```
code du champ = 'valeur recherchée'
```

sachant que chaque propriété est décrite par un champ et un ensemble de valeurs (numériques ou textuelles).

Dans la zone de saisie (Fig. 7) on écrira directement la requête. En reprenant l'exemple précédent où l'on cherche les produits de départ permettant de synthétiser la quinoléine, la saisie sera :

Figure 7 : Arborescence compacte de la base

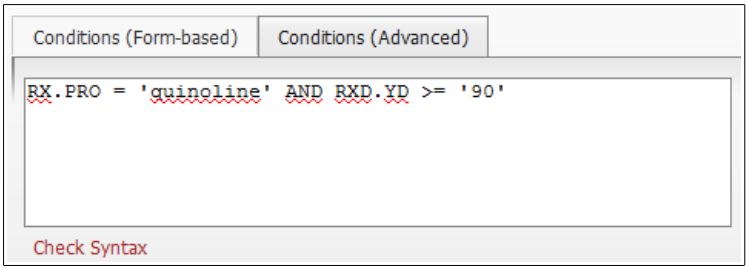


Si l'on ne connaît pas la syntaxe exacte à utiliser, il faudra aller fouiller dans l'arborescence de la base pour retrouver le code du champ. On peut alors soit chercher au "feeling" en développant chaque item soit utiliser l'outil de recherche prévu : **Search for field** (voir entouré).

Remarque : en annexe 2, est donné l'ensemble des champs présents dans la base avec leur code, reflet exact de l'arborescence ci-dessus.

On peut aussi poser une requête contenant plusieurs champs en utilisant les opérateurs booléens AND, OR et NOT. A supposer que l'on s'intéresse aux réactions de préparation de la quinoléine ayant un rendement supérieur ou égal à 90 %, on saisira directement :

Figure 8 :



Cette requête conduit à 11 réactions parmi les 399 possibles sans restriction sur le rendement.

B-II-2- Requête ciblée

On cherche souvent une propriété précise pour une substance donnée. Cette recherche par propriété est dans ce cas tout à fait adaptée; prenons l'exemple supposé d'une recherche du point de fusion du phenanthrene.

On va dans d'abord rechercher le phenanthrene par nom en développant séquentiellement les parties **Identification** puis **Substance Identification** de l'arborescence des champs. On aperçoit sur la Fig. 9a le champ **Chemical Name** activé avec la possibilité de consulter l'index correspondant, symbolisé par 3 petits points (voir entouré).

Remarque : la présence de ces index est systématique pour chaque champ et permet de vérifier si la valeur recherchée est bien présente dans la base.

Figure 9a

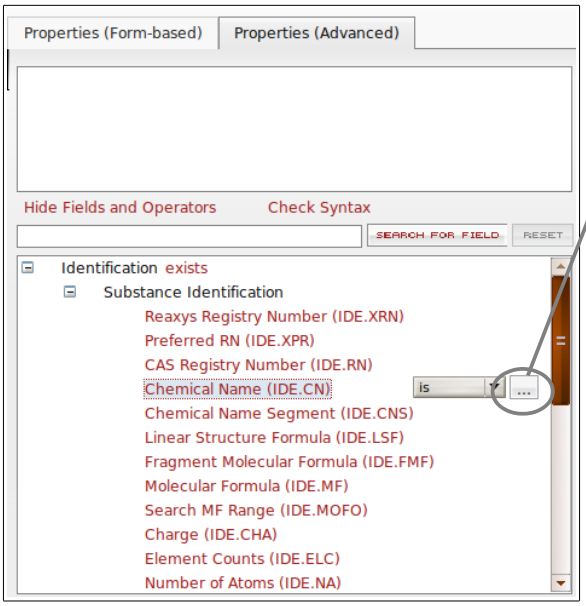
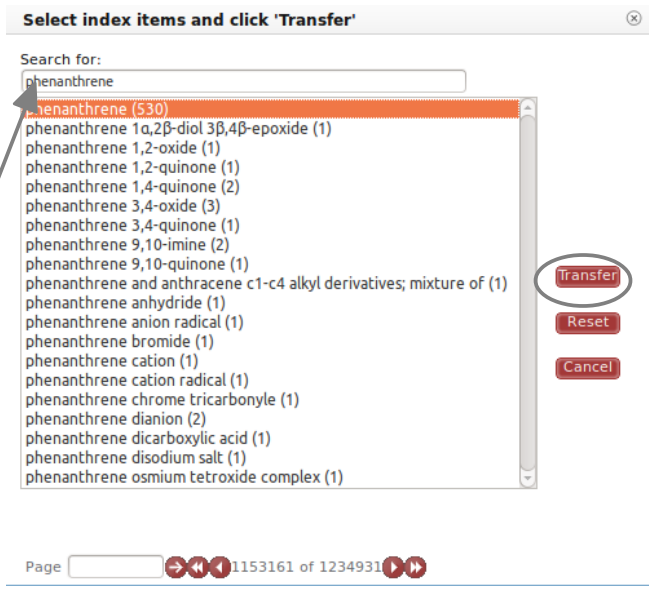
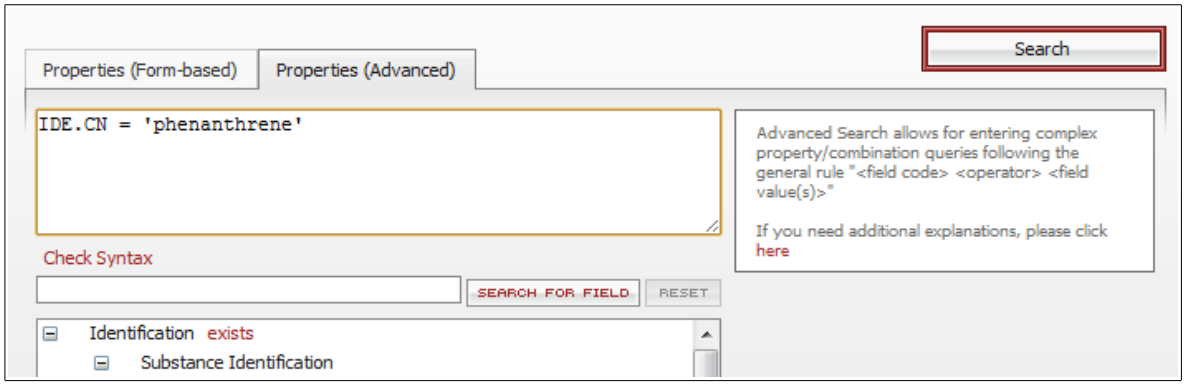


Figure 9b



L'activation de ces 3 petits points conduit à la Fig. 9b ; il suffira d'activer le bouton **Transfer** pour que le terme recherché passe dans la zone de saisie ; l'affichage devient :



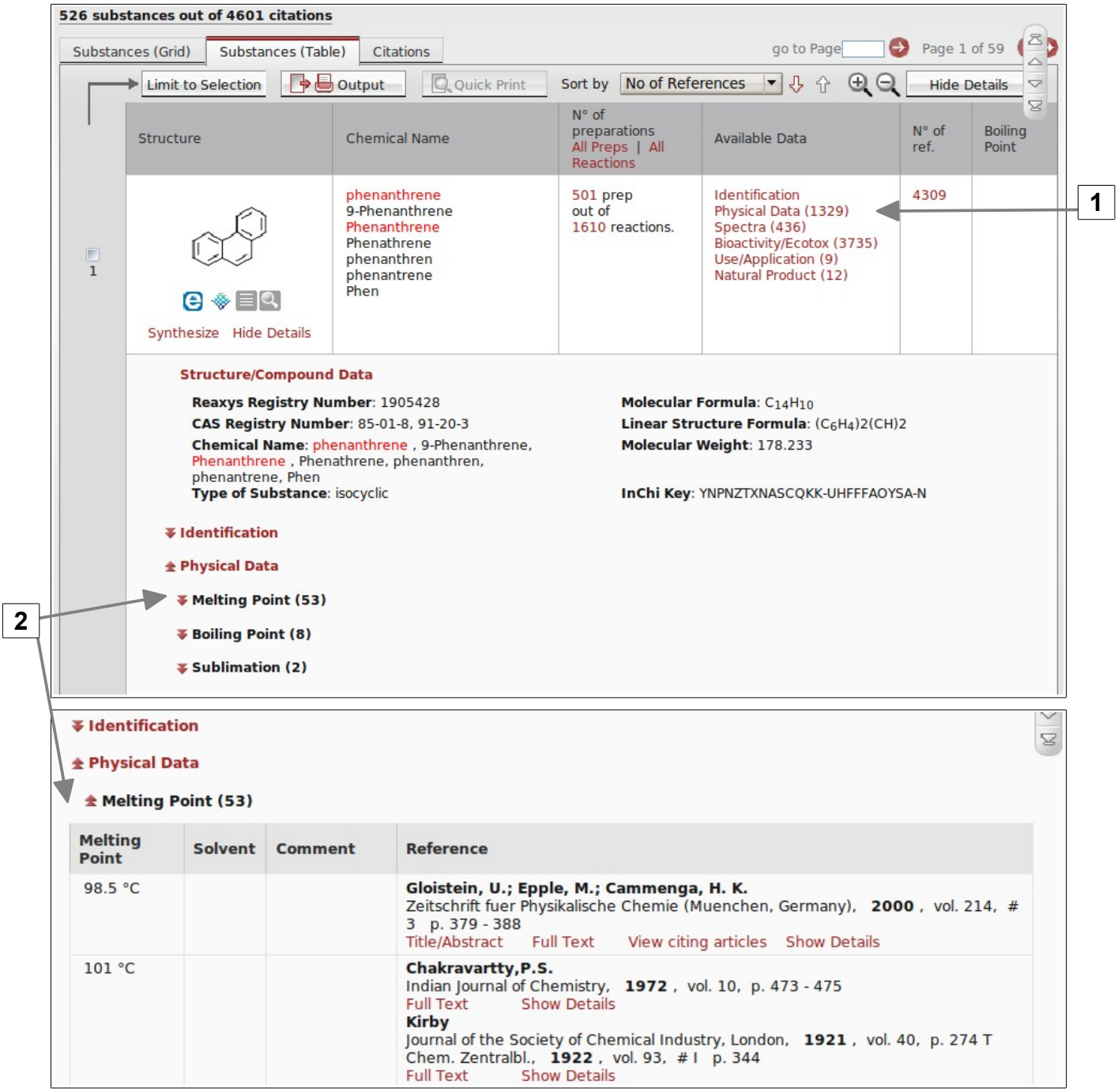
Dans l'exemple choisi, la première substance, parmi les 530 substances trouvées, est exactement le phenanthrene recherché ; c'est celle qui a le plus grand nombre de références ! Il suffira alors d'activer l'hyperlien **Physical data** (voir **1** de la Fig. 10) pour que soit développée cette partie de la zone d'affichage. On aperçoit alors en bas la mention **Melting point** (voir **2** de la Fig. 10) qu'il faudra activer (développer) pour faire apparaître la dernière partie de cette Fig. 10 où sont montrées les deux premières valeurs du point de fusion recherché.

Remarque : dans cet exemple, la recherche par nom chimique a conduit à 526 substances. Pour que la première substance affichée soit la substance recherchée, il suffit de trier la liste de réponses par Nombre de références croissant. Il y a de très grandes chances que ce tri soit pertinent.

Remarque : de manière générale, quand on ne trouve pas une substance en la recherchant par nom, le meilleur moyen est encore de la dessiner (voir section C)dans la zone appropriée (zone ④ de la Fig. 4).

Quel que soit le contexte de la requête (Réactions, Substances ou Citations), la navigation est aisée puisque tous les termes, inscrits en caractères rouge brique, sont des hyperliens soit vers la même page soit pour changer de contexte.

Figure 10 : Exemple de réponse (vue partielle)



B-II-3- Requête ciblée avec troncature

On peut rechercher par mot-clé dans tous les champs textuels mais on peut élargir une requête ciblée en utilisant des troncatures, càd en autorisant des orthographes supplémentaires au terme recherché. Par exemple, supposons rechercher des informations sur des dérivés du discodermolide, produit naturel d'origine marine et possédant une forte activité à la fois immunosuppressive et antifongique.

Si nous utilisons, en mode expert, la commande :

```
IDE.CN = 'discodermolide'
```

qui recherche dans le champ **Chemical Name** la valeur *discodermolide* (cf. B-II-2), nous obtenons 3 substances qui correspondent à la même structure avec des isomères de position.

En plaçant une troncature (*) à droite seulement de la valeur recherchée :

```
IDE.CN = 'discodermolide*'
```

nous obtenons 13 substances dont les 3 précédentes plus 10 dérivés du type : le *3-discodermolide-acétate*, le *3,7-discodermolide-diacétate* ou encore le *3,7,17-discodermolide-triacétate* …

Enfin on peut placer une troncature des deux côtés :

```
IDE.CN = '*discodermolide*'
```

on va élargir le nombre de réponses (64) et retrouver des composés dénommés : *(+)-14-norméthyldiscodermolide*, *C7-épi-discodermolide* ou encore *(-)-discodermolide*.

On peut utiliser les troncatures suivantes (permises à droite, à gauche et au milieu d'un mot) :

- * remplace tous les caractères, simple ou multiple et l'absence de caractère
- ? remplace un caractère simple
- ?? remplace deux caractères.

Pour l'anecdote, la recherche :

```
IDE.CN = 'sul??uric acid'
```

conduit à la réponse ci-dessous, qui illustre les différentes indexations de l'acide sulfurique !

Structure	Chemical Name	N° of preparations All Preps | All Reactions	Available Data	N° of ref.	Boiling Point
 Synthesize Show Details	sulfuric acid conc. H2SO4 H2SO4 sulphuric acid Sulfuric acid acide sulfurique Schwefelsaeure	1038 prep out of 30418 reactions.	Identification Physical Data (594) Spectra (88) Bioactivity/Ecotox (1) Use/Application (216)	26676	

B-II-4- Problèmes d'indexation des réactants, réactifs et catalyseurs

On trouve les définitions suivantes dans le **Help** :

- les **reactants** (réactants, produits de départ) sont des composés qui ajoutent au moins un atome de carbone au squelette au produit d'arrivée. Par ex., MeOH dans la préparation d'un ester ou Ac_2O dans une acétylation. Si un composé agit en tant que produit de départ et que solvant, il est enregistré en tant que reactant.

- les **reagents** (réactifs) sont des composés qui prennent place dans la réaction mais ne contribuent pas au squelette carboné du produit d'arrivée. Par ex., NBS, $AlLiH_4$ sont des **reagents**. Ils sont souvent symbolisés au-dessus de la flèche réactionnelle.

- les **catalysts** sont dans Reaxys catégorisés avec les **reagents**.

Si dans la version précédente (Beilstein Crossfire), les réactants, réactifs et catalyseurs étaient parfois indexés dans un champ ne correspondant pas au rôle exact du composé dans la réaction, les réactants sont *a priori* correctement indexés dans la nouvelle version et bien distincts du groupe des réactifs/catalyseurs.

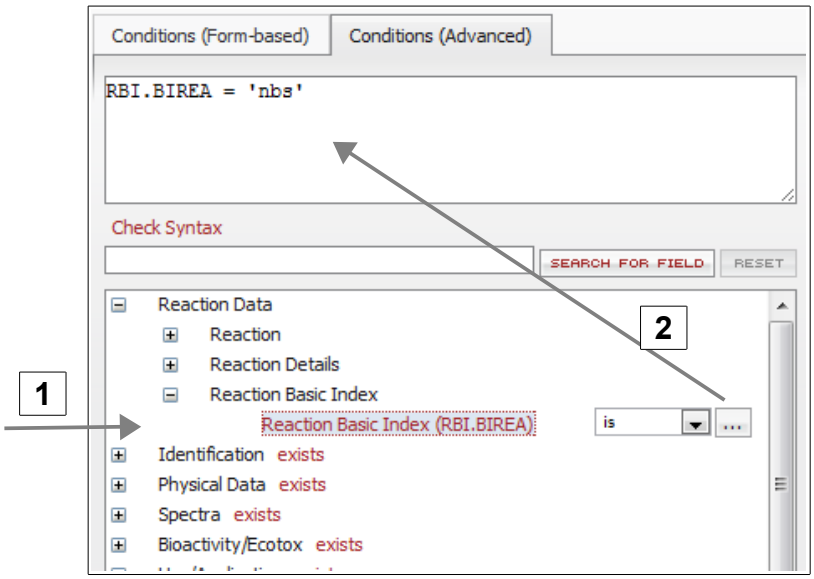
Si l'on prend l'exemple du N-bromo-succinimide (ou NBS), on trouve 1588 réactions où ce composé apparaît comme réactant lorsqu'on utilise les outils de dessin; en le recherchant dans le champ **Reaction Basic Index** , on va trouver 151 268 réactions où le NBS est bien indexé comme un réactif (Fig. 11).

On aurait pu aussi rechercher le NBS dans le champ correspondant à réactif ou catalyseur, mais cette méthode ne fonctionnera pas directement du fait que l'indexation d'une même substance est possible sous plusieurs noms. Pour tenir compte de ces indexations différentes il conviendrait d'utiliser des troncatures et de formuler la question, par ex., sous la forme : *nbs*, *n-brom*succin*, …

L'utilisation du champ **Reaction Basic Index** permet de récupérer toutes les réactions décrivant le NBS comme suit (ce qui tient compte d'indexations possibles en anglais, américain, allemand, français et des fautes d'orthographe dans les 3 langues ….:

```
N-bromosuccinimide / 1-bromo-2,5-pyrrolidinedione / N-bromosuccimide
NBS / N-bromo-succinimide /N-bromosucciniminde
```

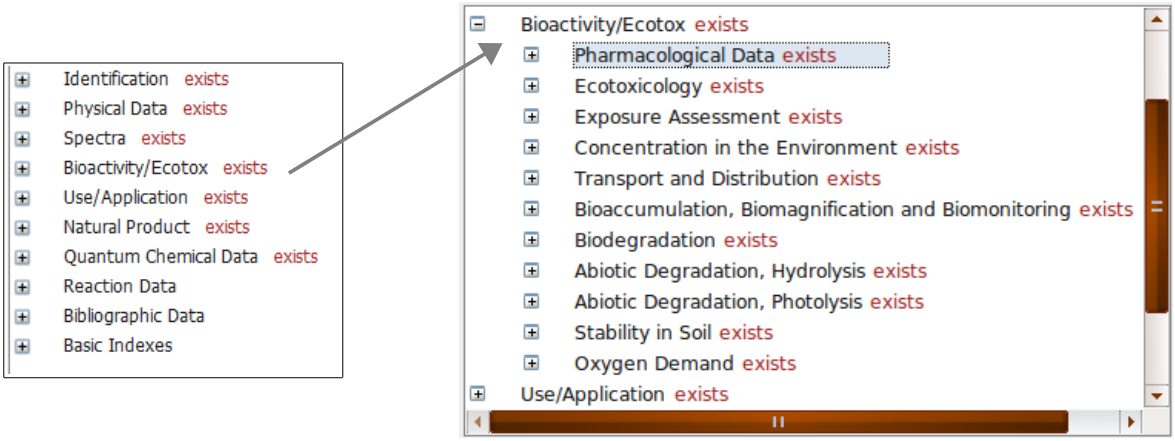
Figure 11 : Recherche d'un réactif et exemple de réponse



B-II-5- Basic Indexes

Ce sont des index collectifs qui évitent d'avoir à prendre en compte individuellement les différents codes des champs (**Field Name**) relatifs à un type de question. Ils donnent donc une vision globale sur un sujet donné et sont aussi intéressants lorsque l'on interroge simultanément par structure ou sous-structure ET avec un **Basic Index**. Voir Requêtes croisées en section F.

On peut se faire une idée de ces index en regardant la Figure ci-dessous puisque la plupart des items se terminent par la mention **exists** et que l'on peut développer chacun de ces items pour faire apparaître de nouveaux index. C'est le cas de l'index **Bioactivity/Ecotox** qui peut être développé en sous-index comme le montre la copie d'écran ci-dessous, sous-index eux aussi développables jusqu'aux champs particuliers.



Très utile pour savoir si un ensemble de substances est décrite, au moins partiellement, par des données spectrales.

Certains anciens (Beilstein) **Basic Indexes** sont toujours disponibles :

Utilisation de	Contexte de requête	Recherche orientée vers
Substance Basic Index (BI.BISUB)	Substances	toutes propriétés des substances
Reaction Basic Index (BI.BIREA)	Réactions	des conditions réactionnelles
Citation Basic Index (BI.BICIT)	Citations	de la bibliographie

Leurs définitions sont précisées en Annexe 1.

B-II-6- Opérateurs logiques (booléens)

On peut interroger simultanément plusieurs champs par des requêtes croisées. Première possibilité : on

utilise les formulaires prédéfinis abordés au paragraphe B-I; deuxième possibilité : on recherche en mode avancé et on doit alors utiliser les opérateurs booléens : **and or, not proximity, near, next** définis ci-dessous. Les opérateurs sont saisis manuellement dans la colonne **Operator** (figure ci-contre) ou bien accessibles à partir du triangle ▼.

L'opérateur AND

permet de retrouver les termes recherchés seulement s'ils sont présents, mais à n'importe quel endroit de l'enregistrement. Exemple :

`MP.MP=1-100 AND exists(NMR)` ==> 633 708 substances ayant un point de fusion entre 0 et 100 °C ET dont un spectre RMN est présent.

L'opérateur OR

permet de retrouver au moins l'un des termes recherchés à n'importe quel endroit de l'enregistrement. Exemple :

`RXD.SOL='ethanol' OR RXD.SOL='chloroform'` ==> plus de 3,8 M réactions réalisées dans l'éthanol OU le chloroforme.

L'opérateur NOT

permet de retrouver des enregistrements vérifiant la première condition et pas la deuxième. Exemple :

`BP.BP=50-60 NOT IDE.ELC = 'F*'` ==> plus de 4,2 M réactions pour des composés ayant un point d'ébullition entre 50 et 60 °C ET ne contenant pas d'atome de fluor.

L'opérateur PROXIMITY

est une variante de l'opérateur AND, pour laquelle les termes recherchés ne sont retrouvés que s'ils sont présents dans une même occurrence d'une propriété. Cet opérateur est très important lorsque l'on recherche des propriétés qui contiennent des sous-champs ou des champs définis par d'autres paramètres. Exemple :

`BP.BP=150 PROXIMITY BP.P=100-150` ==> 121 substances dont le point d'ébullition est égal à 150 °C ET mesuré sous une pression comprise entre 100 et 150 (unité non précisée).

L'opérateur NEAR

est une variante de l'opérateur PROXIMITY, pour laquelle les termes recherchés ne sont retrouvés que s'ils sont présents et adjacents, dans une même occurrence d'une propriété. Exemple :

`AB.AB='fibrinogen*' NEAR AB.AB='receptor*'` ==> 7310 substances recherchées par mots-clés dans le champ 'résumé'.

L'opérateur NEXT

est une variante de l'opérateur NEAR, pour laquelle les termes recherchés ne sont retrouvés que s'ils sont présents et adjacents, dans une même occurrence d'une propriété, et apparaissent dans l'ordre 'premier terme avant le deuxième'. Exemple :

`AB.AB='capillar*' NEXT AB.AB='electrophor*'` ==> 3464 substances recherchées par mots-clés dans le champ 'résumé'.

C – DESSIN DES STRUCTURES ET SOUS-STRUCTURES

C-I- Généralités

La plate-forme Reaxys utilise **MarvinSketch** comme éditeur de structures par défaut; c'est un logiciel libre pour les académiques (https://www.chemaxon.com/) moyennant un enregistrement personnel. Cet éditeur ne nécessite pas d'installation supplémentaire une fois réalisé l'enregistrement à la plate-forme Reaxys. En dehors du cadre des requêtes réalisées avec Reaxys, il est possible d'utiliser MarvinSketch de manière permanente en téléchargeant l'application sur le site de ChemAxon après s'être enregistré comme utilisateur.

Remarque : MarvinSketch fonctionne avec une "machine" Java (version 1.5 ou plus récente). La plupart des utilisateurs travaillant avec ScifinderWeb n'auront pas d'installation supplémentaire à réaliser. Si besoin, il suffit de se connecter sur le site de Sun pour installer la machine Java appropriée à son système d'exploitation (http://www.java.com/fr/download/manual.jsp).

C-I-1- Editeurs de structure externes

Reaxys peut être configuré pour fonctionner avec les éditeurs externes, listés ci-dessous, sous les différentes versions de Windows ou Linux, mais pas pour Macintosh. Des infos plus détaillées sont disponibles dans le **Help** de l'interface de Reaxys.

- ***CrossFire Structure Editor***

Le **CrossFire Structure Editor** est l'éditeur historique utilisé par défaut dans la version du logiciel **CrossFire Commander**. Cet éditeur permet l'accès aux mêmes données, structures et réactions que l'on peut trouver dans la base Reaxys. Il doit être installé sur chaque ordinateur avant d'utiliser Reaxys (info.crossfirecommander.com).

- ***Symyx ISIS Draw***

Symyx MDL® ISIS/Draw est l'éditeur habituellement employé avec le logiciel ISIS/Base. Il nécessite lui-aussi une installation préalable sur le poste d'interrogation, sachant qu'il est gratuit pour les académiques (www.Symyx.com).

- ***Symyx Draw***

Symyx MDL® Draw est un logiciel similaire au précédent fonctionnant avec une machine virtuelle Java; il ne sera pas présenté dans les paragraphes suivants (www.symyx.com).

- ***CambridgeSoft ChemDraw***

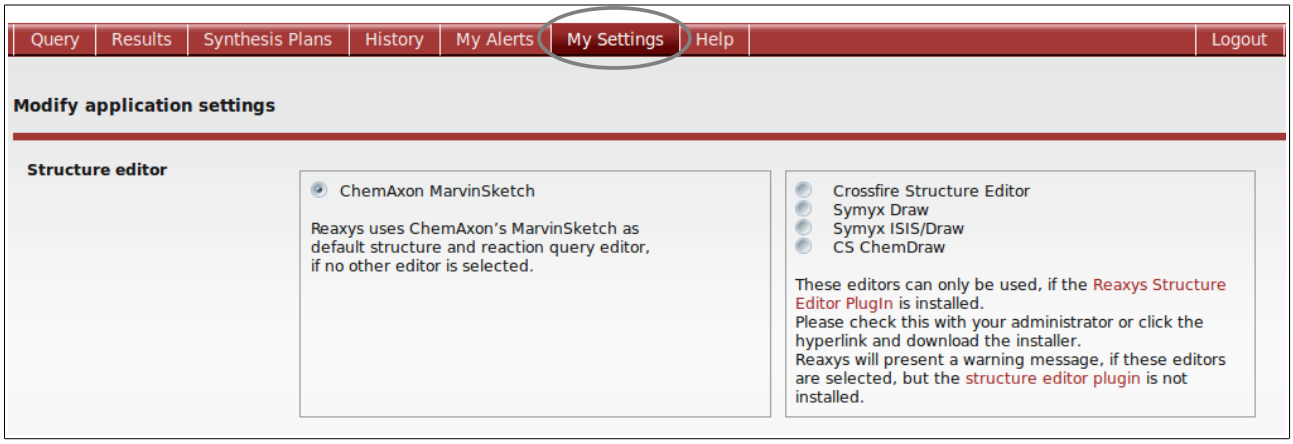
ChemDraw est l'éditeur de structures développé par la société **CambridgeSoft ChemDraw**. Il est particulièrement utile en raison de ses très nombreux **templates** structuraux prédéfinis et ses modèles de style de publication dans les journaux liés à la chimie (http://scistore.cambridgesoft.com/).

- ***InfoChem ICEdit***

ICEDIT est l'éditeur développé par la société InfoChem Gmbh; (http://www.infochem.de/products/software/icedit.shtml).

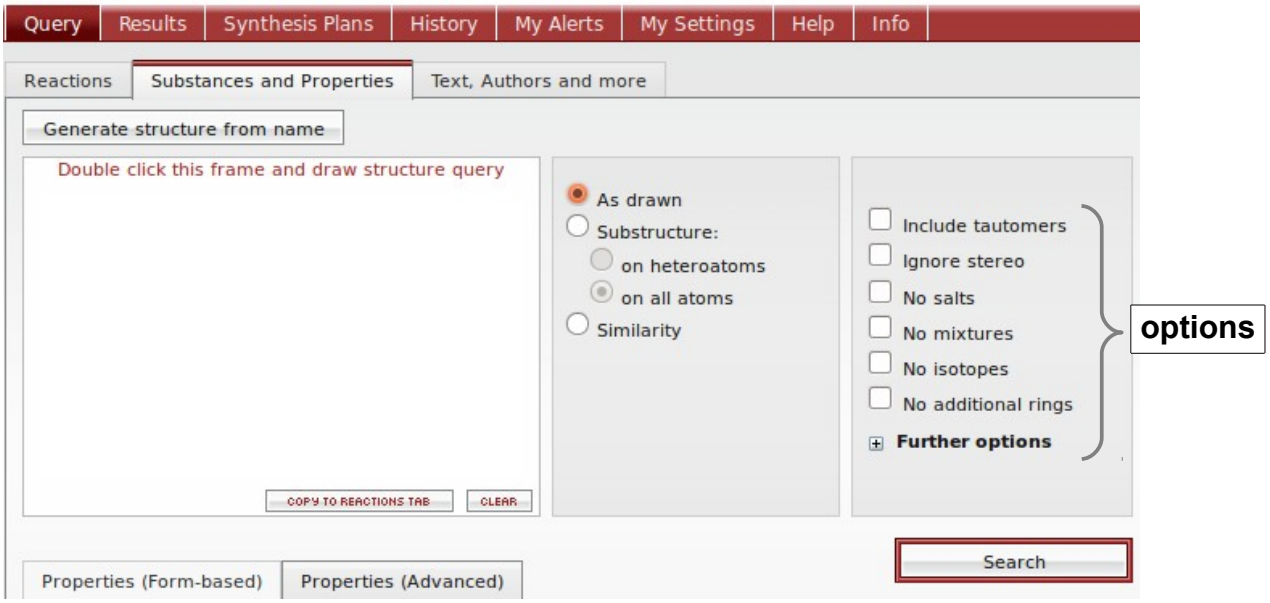
Remarque : quel que soit l'éditeur externe utilisé, il faudra installer un plug-in supplémentaire pour pouvoir dessiner des molécules que l'on pourra se fournir à l'adresse suivante : http://www.info.reaxys.com/support_downloads. Avant d'installer le plug-in approprié, il faudra s'assurer que le logiciel correspondant est lui aussi correctement installé.

Une fois ces installations réalisées, la personnalisation de l'éditeur de son choix se fait dans Reaxys comme suit. Dans le bandeau principal de l'application, on active **My Settings** ce qui ouvre la boîte de dialogue ci-dessous, où l'on pourra confirmer son choix.


Query
Results
Synthesis Plans
History
My Alerts
My Settings
Help
Logout
Modify application settings
Structure editor
ChemAxon MarvinSketch
Reaxys uses ChemAxon's MarvinSketch as default structure and reaction query editor, if no other editor is selected.
Crossfire Structure Editor
Symyx Draw
Symyx ISIS/Draw
CS ChemDraw
These editors can only be used, if the Reaxys Structure Editor PlugIn is installed.
Please check this with your administrator or click the hyperlink and download the installer.
Reaxys will present a warning message, if these editors are selected, but the structure editor plugin is not installed.

Nous allons voir dans les paragraphes suivants les outils de l'éditeur de structure **Marvin Sketch** et ses interactions avec l'interface de requête. Si nous considérons la Fig. 12, c'est donc à partir la zone centrale (**Double click this frame and draw structure query**) que l'on pourra ouvrir l'éditeur de structures choisi. Nous aborderons les options de requête qui figurent à droite de l'interface au chapitre suivant.

Figure 12



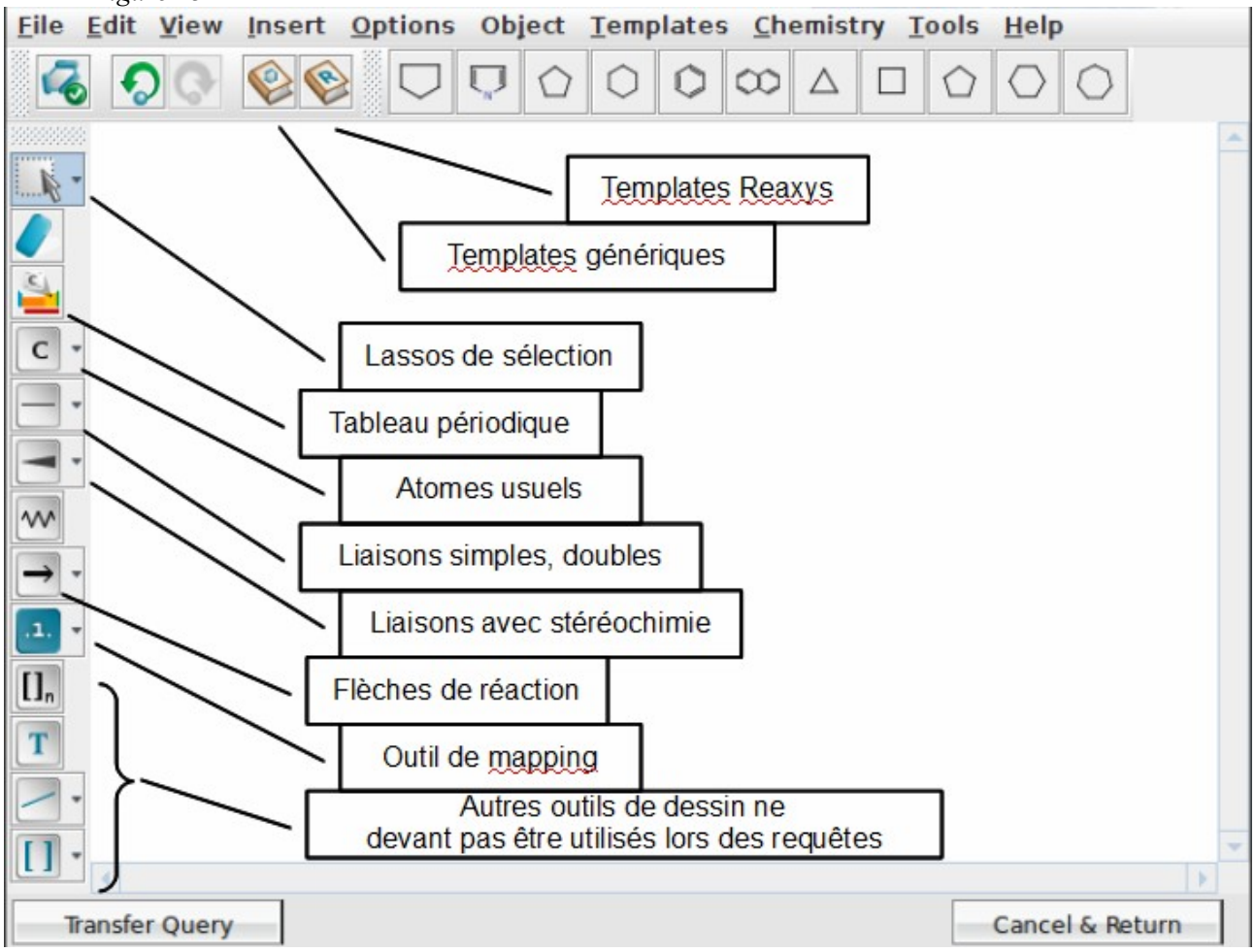
Remarque : dans le cadre de ce manuel, nous décrirons les outils proposés par MarvinSketch. Les utilisateurs habitués des autres éditeurs pourront retrouver la description des outils décrits dans les paragraphes suivants en consultant le ***Help*** *de l'interface, à la rubrique* ***Structure/Reactions Editors****.*

C-II- Dessiner avec MarvinSketch

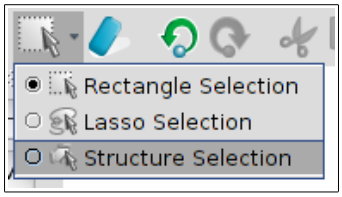
C-II-1- Interface de dessin

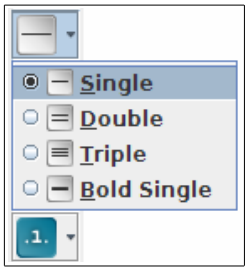
MarvinSketch est équipé d'outils de dessin dont les principaux sont résumés sur la Fig. 13 et que nous allons détailler dans les paragraphes suivants.

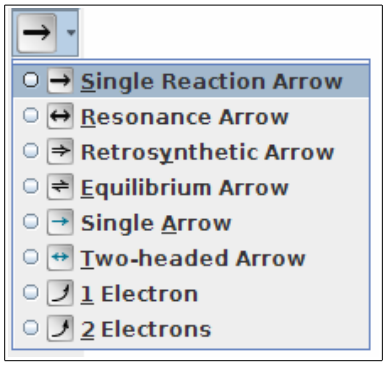
Figure 13



Chaque outil de la barre verticale possédant un petit triangle noir présente plusieurs options que l'on peut faire apparaître en maintenant le clic gauche enfoncé quelques secondes sur l'icône correspondante (voir Figures ci-contre).

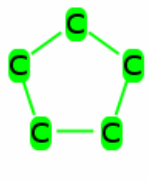
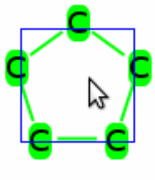
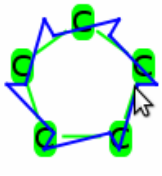






Déplacer – Tourner – Zoomer une structure

Lorsque l'on doit déplacer une structure globalement, il faut d'abord la sélectionner (Fig. 14a) puis positionner le curseur vers le centre de la structure jusqu'à apparition du rectangle puis déplacer en maintenant le clic gauche actif (Fig.14b). Pour faire tourner une structure, il faut la sélectionner (Fig. 14a) puis positionner le curseur à la périphérie de la structure jusqu'à apparition de la roue dentée et faire tourner en maintenant le clic gauche actif (Fig. 14c)

Fig. 14a	Fig. 14b	Fig. 14c
		

Si l'on veut zoomer sur la feuille de dessin, il faut activer la commande du menu **View/Transform/Zoom** (touche F6) puis déplacer la double flèche, apparue en remplacement du curseur, de bas en haut.

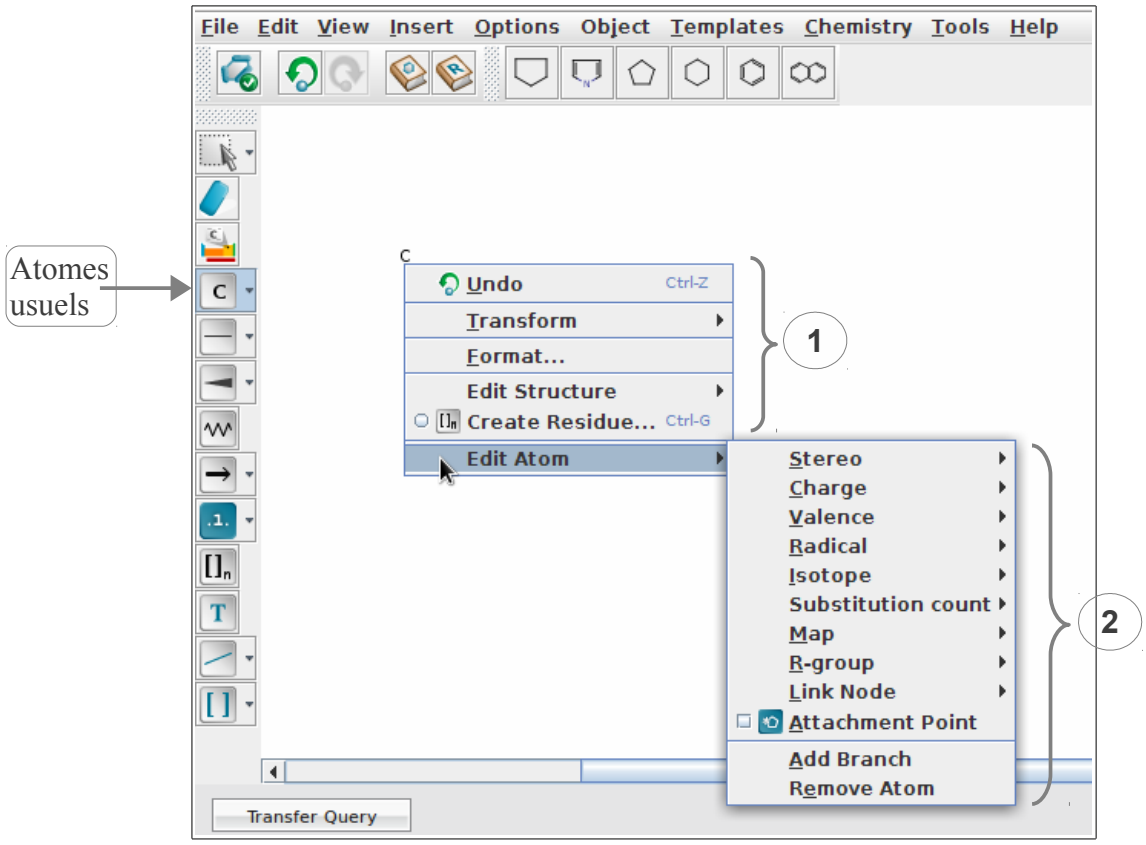
C-III- Atomes

C-III-1- Spécification des atomes

Pour introduire un atome (C dans cet exemple), on dispose de l'outil Atomes usuels de la Fig. 13. L'atome étant activé, on peut éditer ses attributs par un clic-droit ce qui conduit au menu déroulant ① puis en bas de ce menu, par la commande **Edit atom**, au menu déroulant ②.

Remarque : ce menu déroulant est accessible par la commande du menu Object/Atom.

Figure 15



La liste des attributs d'un atome est importante puisque l'on peut modifier :

- la stéréochimie (concerne les modifications de la stéréochimie au cours d'une réaction)
- la charge,
- la valence (par ex. celle de l'atome d'azote peut être 4 ou 5 pour les groupements ammonium et nitro respectivement).
- la nature d'un radical (monovalent, divalent …)

- éditer un isotope non naturel
- changer le nombre de substitutions de cet atome (détaillé ci-dessous)
- éditer le mapping de l'atome au cours d'une réaction (cf. E-III-1)
- introduire un groupement variable (R-group – cf C-VIII)

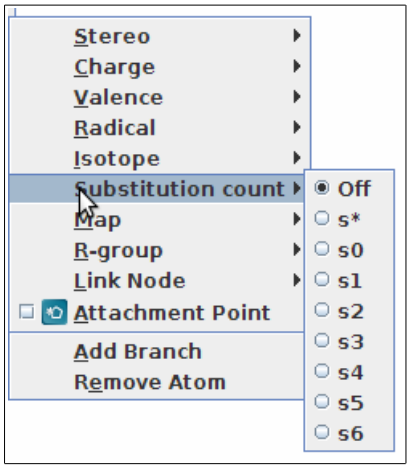
C-III-1- Substitution des atomes

L'intérêt des recherches sous-structurales est de pouvoir spécifier que tel ou tel atome de la sous-structure est substituable ou non.

Remarque : par défaut une structure (ou sous-structure) est recherchée sans substituant, donc avec des H par défaut si rien n'est précisé pour le statut des atomes.

Si on reprend l'exemple de la Fig. 15 et que l'on développe les menus déroulants jusqu'à **Substitution Count,** on voit apparaître plusieurs options permettant de fixer le nombre de substituants de l'atome concerné :

- **off**, valeur par défaut
- **s*** empêche tout autre substituant sur cette position.
- **s0** précise que l'atome ne comporte aucun substituant
- **s1** précise que l'atome comporte un seul substituant
- ….
- **s6** précise que l'atome peut comporter jusqu'à 6 substituants.

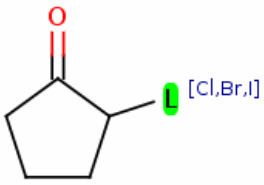


Astuce : pour définir le degré de substitution de plusieurs atomes, avec s4 par exemple, on sélectionne le premier atome puis on saisit au clavier "Majuscule-point" ce qui fait apparaître une liste d'options; en relâchant le clavier on saisit ensuite "s4", option qui reste active tant qu'on ne change pas d'outil.

C-III-2- Liste d'atomes – Exclusion d'atomes

Pour définir une (ou plusieurs) liste(s) d'atomes, on sélectionne l'atome concerné, puis on choisit l'outil "Tableau périodique" de la Fig. 13 ce qui ouvre la boîte de dialogue de la Fig. 16. Dans cette boîte, on active **Atom list** et il suffit de cocher les atomes souhaités – Cl, Br et I dans cet exemple.

On obtient par exemple le dessin ci-dessous :



De la même manière, on peut exclure une liste d'atomes avec l'option **Not List** de cette même boîte.

C-IV- Liaisons

A partir d'une liaison dessinée et sélectionnée, on va pouvoir ouvrir par un clic-droit les menus déroulants de la Fig. 17 où, par la commande **Edit bond** en bas du premier menu, on accède à un deuxième menu déroulant.

Remarque : ce menu déroulant est accessible par la commande du menu Object/Bond.

Figure 16

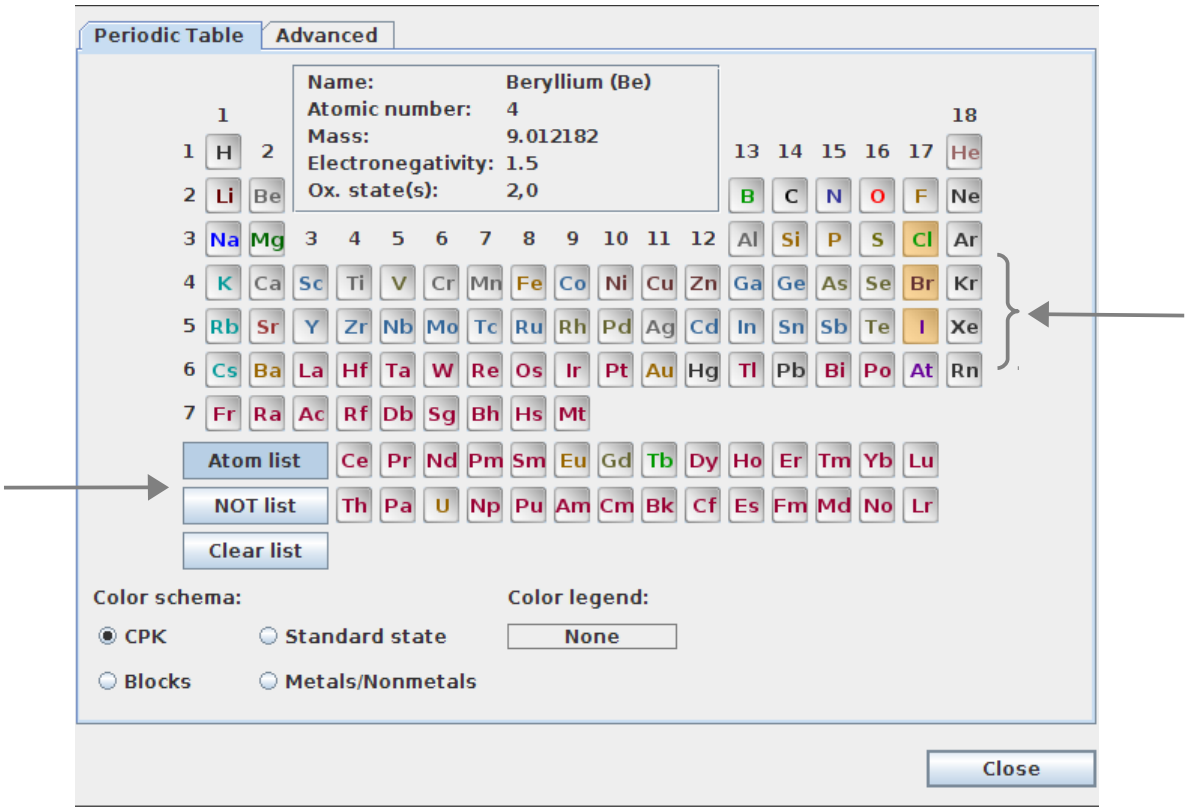
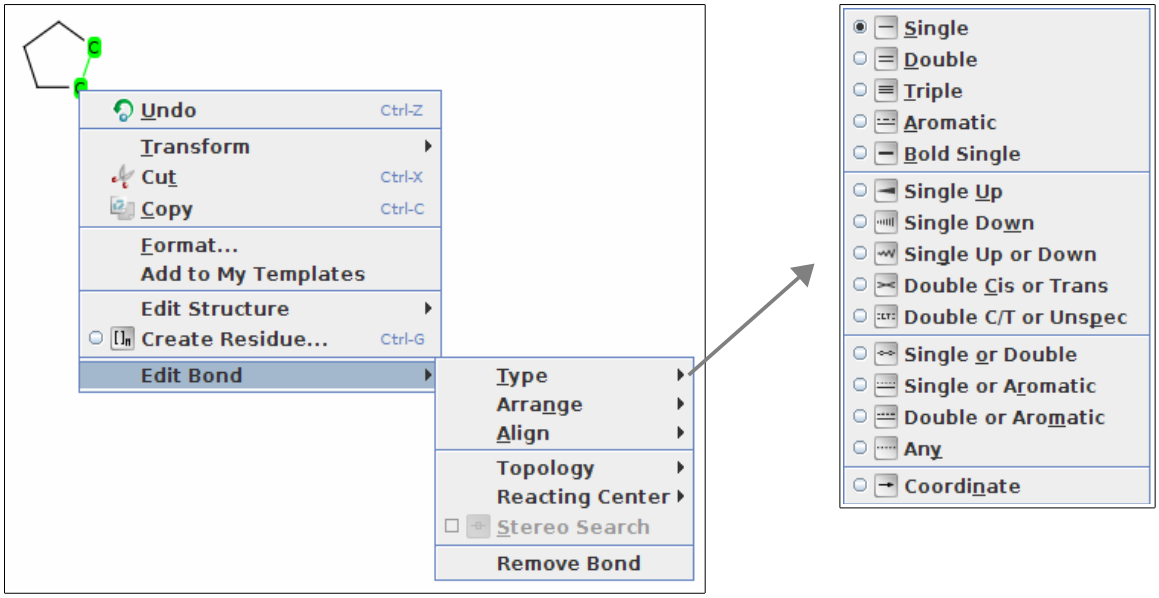


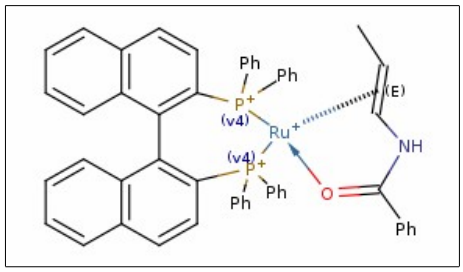
Figure 17



Les principaux attributs d'une liaison sont : son type (simple, double, triple, attributs stéréochimiques, …), sa topologie (cycle ou chaîne ou indifférent), définir si cette liaison est centre de réaction ou non (**Reacting Center,** cf. E-III-2).

C-IV-1- Composés de coordination

Supposons que l'on veuille dessiner le composé de coordination suivant, intermédiaire réactionnel supposé lors d'une hydrogénation, où l'atome de Ruthénium possède une liaison dative σ avec l'atome d'oxygène de l'amide et une liaison de coordination π avec la double liaison à hydrogéner.



Etape 1 : on commence par dessiner le BINAP auquel on ajoute l'atome de Ruthénium; on note au passage que la valence des deux atomes de phosphore a été fixée à 4 (Fig. 18a). De même la charge de l'atome de Ruthénium est portée à +1.

Etape 2 : on construit la liaison dative entre l'atome d'oxygène en dessinant d'abord une liaison simple de l'Oxygène vers le Ruthénium puis on change son type (cf. Fig. 17) pour le type **Coordinate**.

Etape 3 : on construit la liaison de coordination entre la double liaison et le Ruthénium. Pour ce faire, on sélectionne la double liaison et par un clic-droit on active successivement les commandes **Edit Structure/Add/Multi-Center** (Fig. 18b). Il apparaît alors une petite étoile au-niveau de la double liaison. Enfin on crée une liaison simple à partir du centre de la double liaison vers le Ruthénium dont on fixe le type à **Coordinate** comme à l'étape 2.

On obtient la figure souhaitée au départ et on notera au passage la différence de représentation entre les deux liaisons de coordination qui ont pourtant été créées de la même manière.

Figure 18a

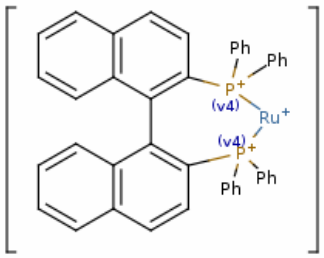
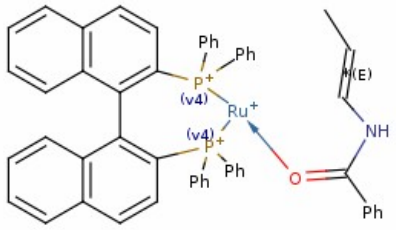
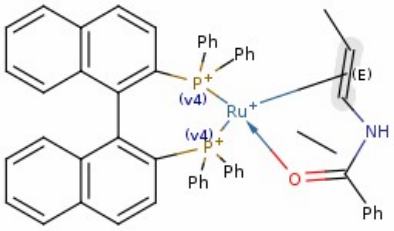
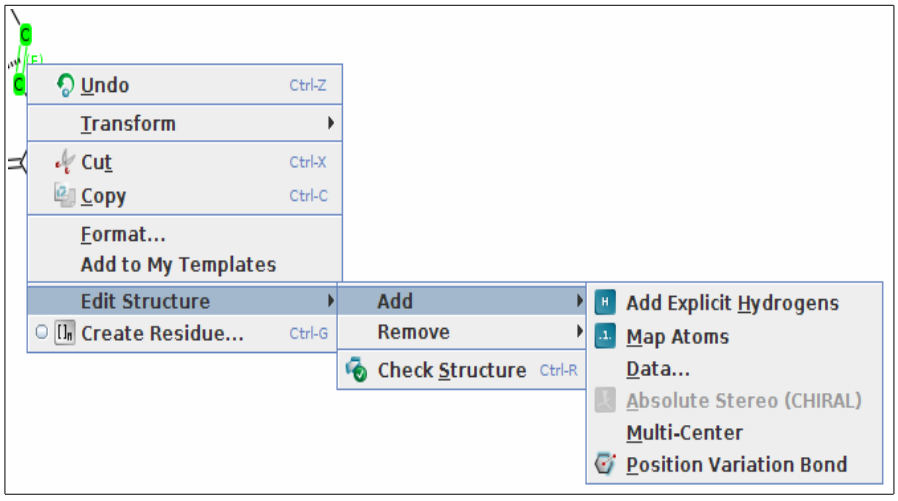
Etape 1	Etape 2	Etape 3
		

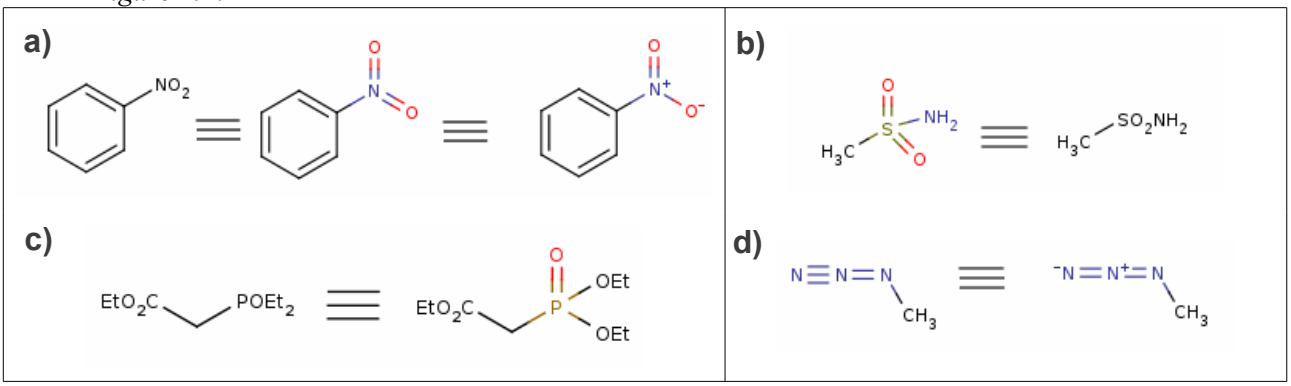
Figure 18b



C-V- Groupements Fonctionnels

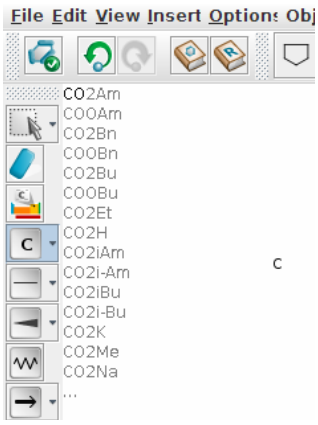
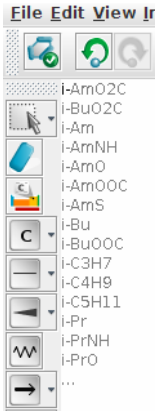
Les groupements fonctionnels peuvent être dessinés liaison après liaison mais aussi de manière intuitive comme pour le nitrobenzène (Fig. 19a); on remarquera au passage que les 3 représentations du nitrobenzène sont valides lors d'une requête car elles sont reconnues comme équivalentes. Il en est de même pour les groupements sulfonamide (Fig. 19b), diéthylphosphonate (Fig. 19c) ou encore azoture (Fig. 19d).

Figure 19 :



Par conséquent pour dessiner rapidement des groupements fonctionnels tels que NO2, CO2Et, OAc … on peut les saisir directement au clavier puisque de toute façon ces groupements sont reconnus lors des requêtes.

En pratique on sélectionne l'atome qui doit devenir le futur groupement fonctionnel – on peut aussi démarrer à partir d'une feuille de dessin vierge – et on saisit le premier ou les deux premiers caractères appropriés; un affichage temporaire apparaît en haut à gauche de la "feuille de dessin" et on peut continuer sa saisie en fonction des valeurs affichées. Quelques exemples de ces affichages temporaires sont donnés sur les copies d'écran ci-dessous; si la saisie commence par ...

… CO, alors on obtient	NH	NO	O	i-
				

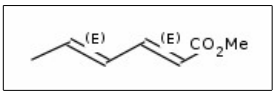
C-VI- Stéréochimie

Pour les doubles liaisons, lorsque l'on dessine la molécule avec la stéréochimie souhaitée, les attributs (*E*) ou (*Z*) s'affichent automatiquement en fonction de la stéréochimie dessinée (Fig. 20a).

Si l'on veut effectuer une recherche avec une double liaison de stéréochimie spécifiée Z ou E, il faudra : 1) le spécifier dans MarvinSketch en cochant la case **Stereo Search** comme indiqué sur la Fig. 20b, 2) le préciser dans l'interface de requête en laissant la case **Ignore Stereo** non cochée (cf. options de la Fig. 12).

Figure 20

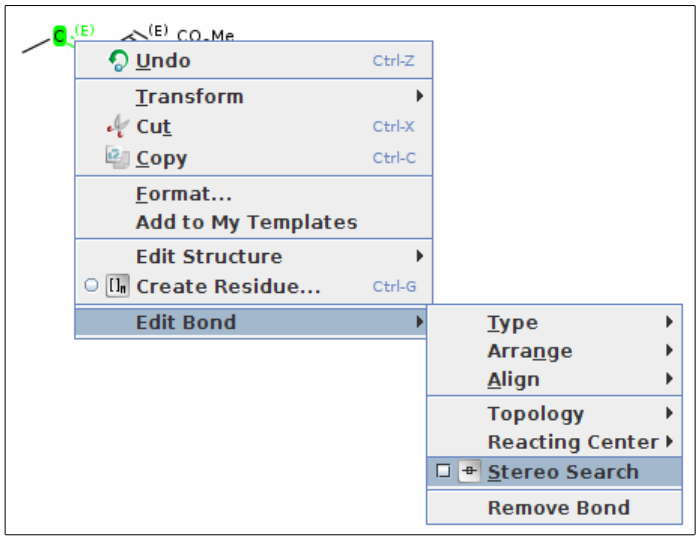
a)



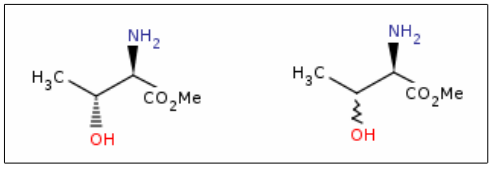
Affichage par défaut de la stéréochimie des doubles liaisons du sorbate de méthyle.

Remarque : Dans l'interface de requête on a laissé la case ***Ignore Stereo*** *non cochée (cf. options de la Fig. 12).*

b)



Pour les liaisons simples elle est précisée, selon la méthode habituelle, par le dessin de la liaison par **Up** ou **Down** (cf. Fig. 13 et/ou 17). L'exemple ci-contre illustre deux recherches qui conduisent dans le premier cas à une seule réponse – l'ester méthylique de la (2R,3R)-thréonine – dans le deuxième cas à 3 réponses – la précédente, l'ester méthylique de la (2R,3S)-thréonine et l'ester méthylique de la thréonine où seule la stéréochimie du C2 est de configuration connue (2R).



Remarque : les notions de stéréochimie absolue ou relative ne sont pas prises en compte lors des requêtes.

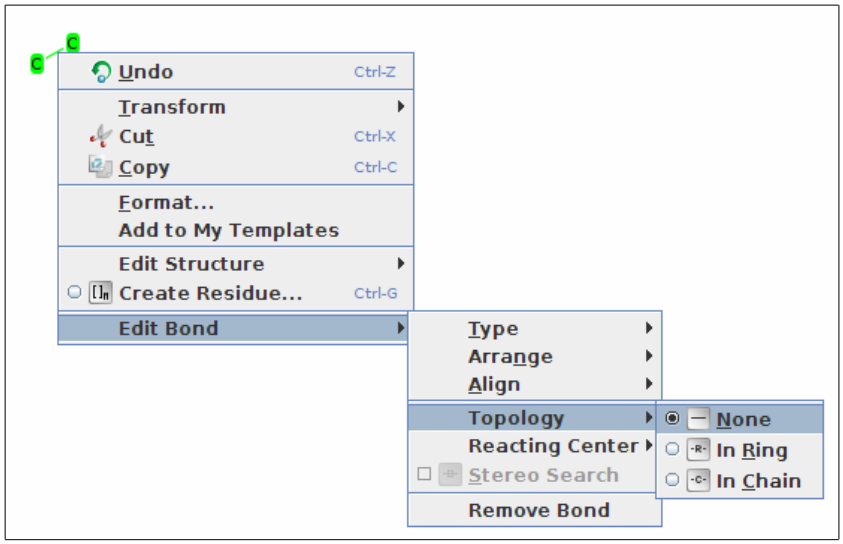
C-VII- Topologie

C-VII-1- Généralités

La topologie se précise au niveau des liaisons en imposant qu'une liaison doit se trouver :

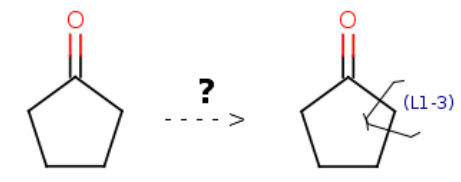
- dans une chaîne ou dans un cycle (option par défaut)
- dans un cycle
- dans une chaîne comme indiqué sur la Fig. 21 (menu déroulant accessible par un clic-droit).

Figure 21

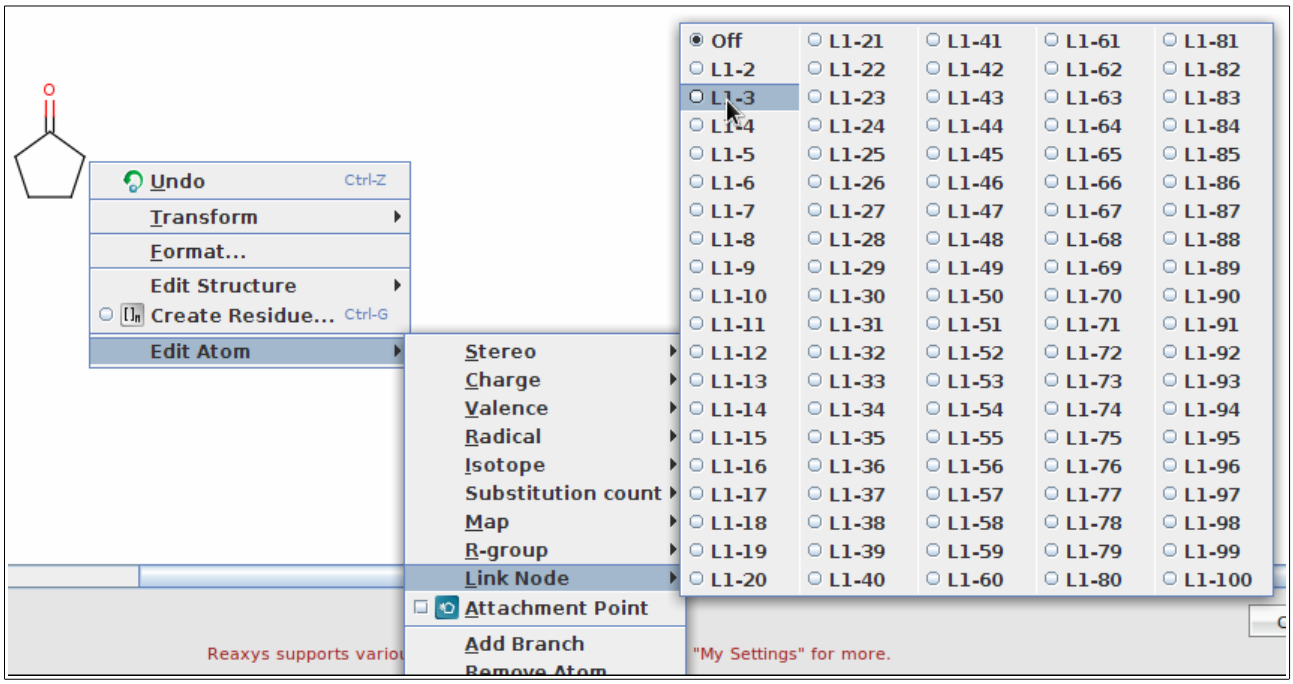


C-VII-2- Chaînes ou cycles de taille variable

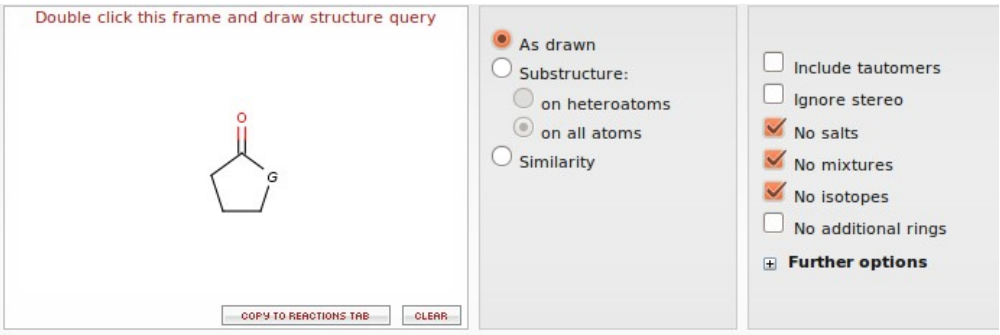
Pour spécifier qu'une chaîne ou un cycle soit de taille variable, on pourra utiliser l'outil **Link Node** comme suit. Supposons que l'on veuille chercher des cétones cycliques dont la taille varie de 5 à 7 ...



… on effectue un clic-droit en α de la fonction cétone par exemple et l'on déroule successivement les items **Edit Atom**, **Link Node** et **L1-3** dans cet exemple.



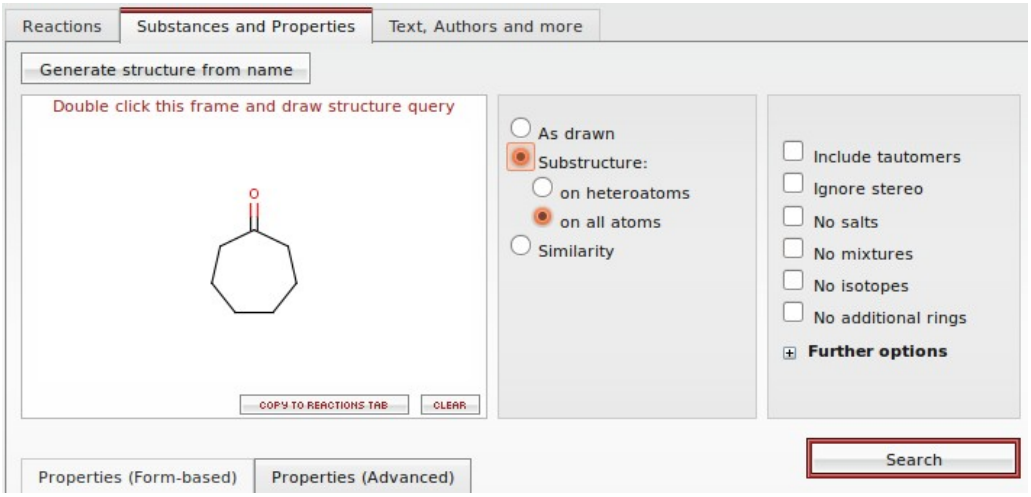
Remarque : on pourrait élargir cette question à des hétérocycles de taille ≥ 5 en introduisant à la place du Link Node le groupement prédéfini G (cf. C-VIII-1); on obtiendrait toutes sortes de cycles de taille ≥ 5, y compris avec un ou des hétéroatome(s). La question, telle que posée ci-dessous, fournit 103777 substances.



C-VII-3- Cycles (condensés, pontés, spiro)

Dans les options de recherche (Fig. 22), on peut activer ou désactiver l'option **No additional rings**, pour que soient inclus ou non les structures possédant, sur les atomes de la sous-structure, des cycles autres que ceux qui ont été explicitement dessinés.

Figure 22



Si l'on pose la question telle que ci-dessus y compris les options, on obtient 36 850 réponses concernant toutes sortes de monocycles à 7, mais aussi des polycycles – condensés, pontés, spiro – contenant la cycloheptanone recherchée dont certains ne seront pas directement rattachés au "cycle-requête", comme sur la molécule ci-contre.

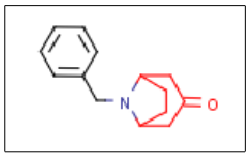
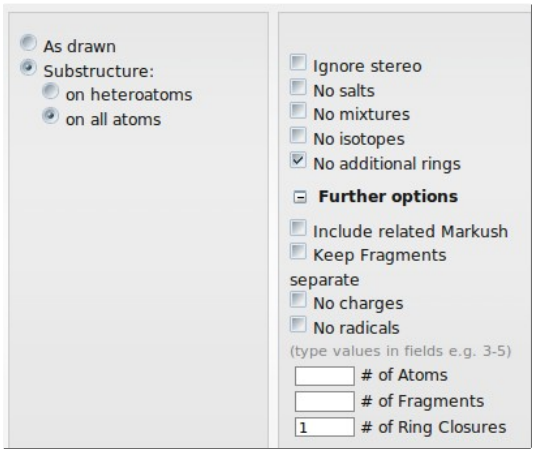


Figure 23

Pour n'obtenir que des monocycles à 7, il faudra poser la même question en ayant coché l'option **No additionnal rings**. Ensuite dans l'interface de requête, on développe la partie **Further Options** et on inscrit la valeur 1 pour l'item **# of Ring Closures**. On obtient 1946 monocycles à 7.



Comment obtenir sélectivement des cycles condensés, pontés, spiro ?

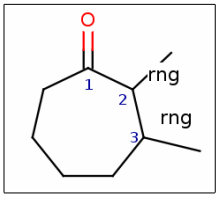
On reprend l'exemple de la cycloheptanone à laquelle on veut accoler un cycle de taille variable en positions 2 et 3. On commence par dessiner deux liaisons exocycliques sur ces deux positions et on leur attribue le statut de "liaison de cycle" comme indiqué sur la Fig. 21. On aboutit au dessin de la Fig. 24a.

Pour la requête, on fixe la valeur 2 (bicycles) à l'item **# of Ring Closures** (cf . Fig. 23) ce qui conduit à 538 bicycles correspondant à la requête (Attention à décocher l'option **No additionnal rings**).

Figure 24

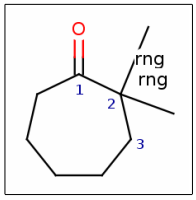
a)

bicycles pontés



b)

bicycles spiro



Pour un bicycle spiro, on utilisera une requête analogue en suivant le dessin de la Fig. 24b.

Dans le cas de ces deux requêtes, les deux atomes exocycliques sont des atomes de Carbone ce qui empêche la présence d'hétéroatomes dans ces deux positions.

C-VIII- Groupements variables

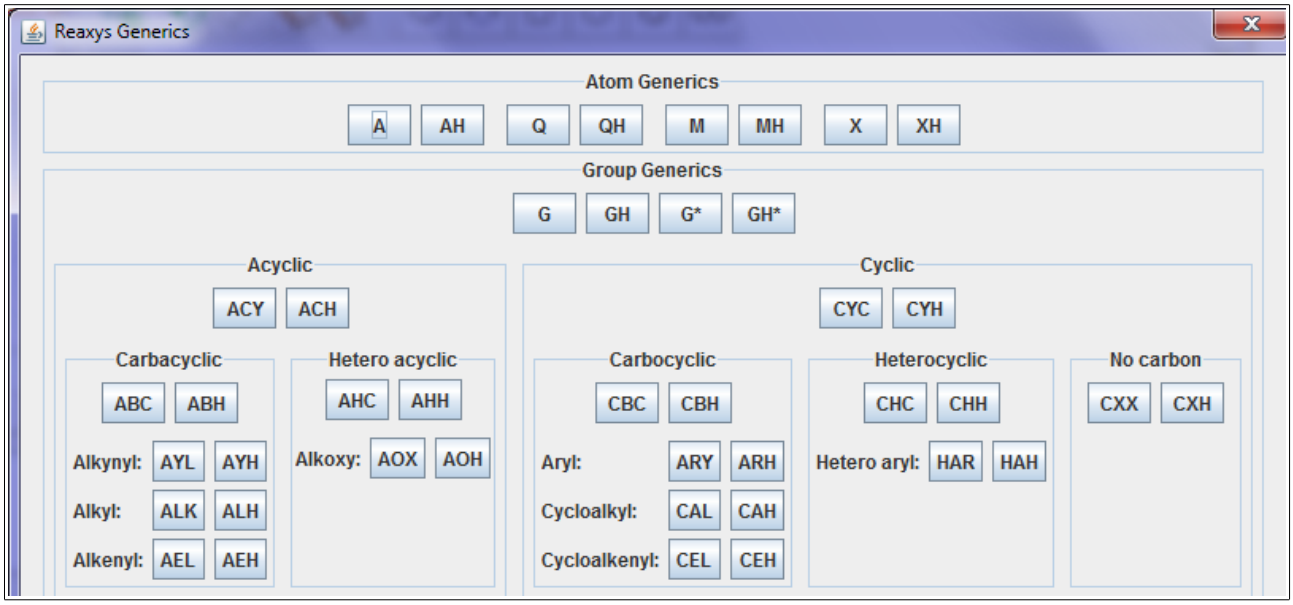
On peut dessiner des structures ayant des groupements variables soit prédéfinis (alkyle, aryle, hétérocyclique, ...) soit définis par l'utilisateur.

C-VIII-1- Groupements prédéfinis

Ils sont regroupés sous la dénomination **Reaxys Generic** et accessibles *via* l'icône de la barre d'outils de MarvinSketch montrée ci-dessous.

reaxys.com
https://www.reaxys.com/reaxys/js/sre_2_1_0_5/child_java.jsp
File Edit View Insert Options Object Templates Chemistry Tools Help

Figure 25 : Groupements génériques prédéfinis



- On ne peut utiliser ces groupements prédéfinis que liés à un atome, réel ou générique.
- La différence entre G et GH vient de ce que dans le deuxième cas le proton est admis dans la requête.
- La différence entre G et G* provient de ce que ce dernier englobe les cycles dans la requête.

C-VIII-2- Groupements variables et structures de Markush

Ces groupements variables, définis par l'utilisateur, permettent de définir une classe de composés grâce à une structure unique appelée structure de Markush qui décrit cette classe par des notations génériques. Cette notation est souvent rencontrée dans les revendications des brevets et la description des librairies combinatoires.

Dans Reaxys ils sont regroupés sous le terme **RGroup** et accessibles par la liste déroulante de la Figure 26. Nous prendrons deux exemples, le premier concernant le cas d'une sous-structure avec un seul point d'attachement, le deuxième avec deux points d'attachement.

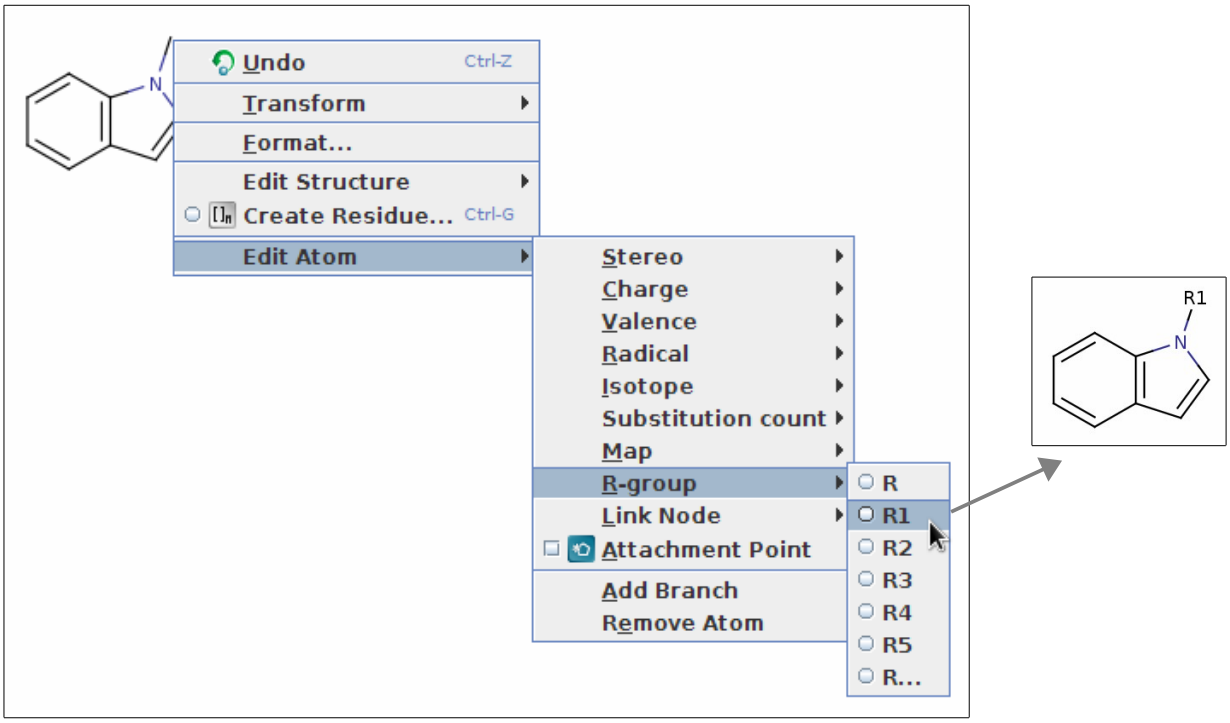
C-VIII-2-a- Un seul point d'attachement

Nous chercherons l'ensemble des indoles N-substitués par un groupement méthyle, éthyle, isopropyle ou acide carboxylique. La procédure décrite ici fait intervenir 3 étapes A, B et C:

A - Positionnement du groupement variable

La sous-structure indolique étant dessinée, on active l'atome à la place duquel doit se trouver le groupement variable à définir. On choisit *via* les menus déroulants la commande **Edit Atom/RGroup/R1** (cf. Fig.26). On peut définir plusieurs groupements variables sur la même sous-structure.

Figure 26

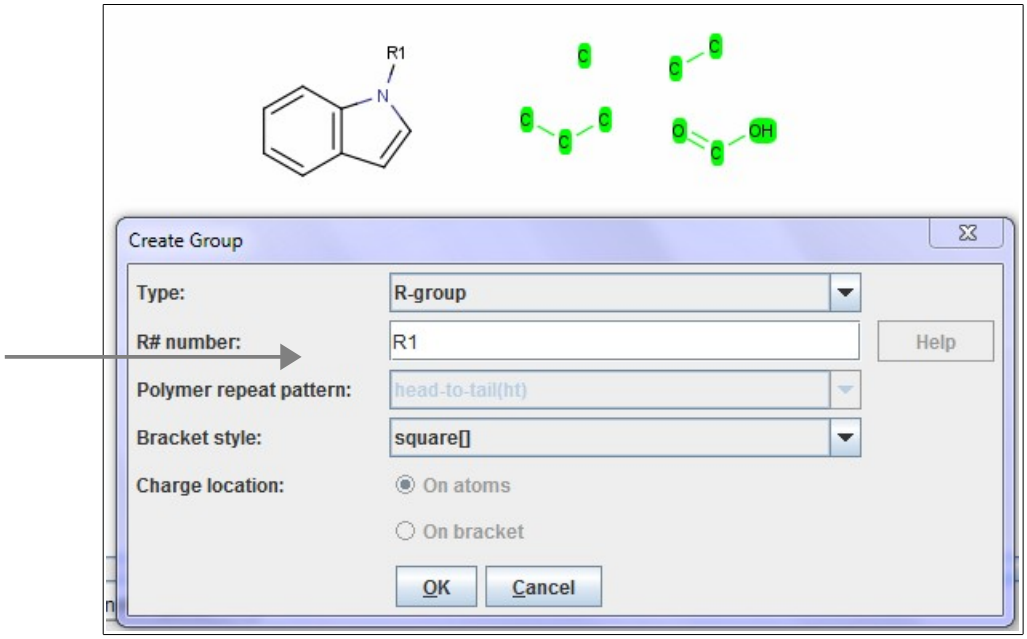


B - Dessin des sous-structures

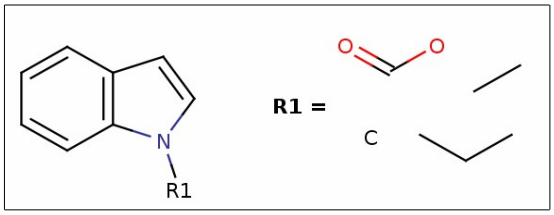
On dessine alors les fragments souhaités, on les sélectionne pour les inclure dans le groupement R1, puis clic-droit et on active la commande **Create Residue** (Fig. 27). Au-niveau du premier item de la boîte de dialogue résultante on choisit **R-group**, puis R1 au-niveau du deuxième item **R# number**.

Astuce : une fois les fragments sélectionnés, on arrive au même résultat en tapant R1 au clavier.

Figure 27



Ceci conduit à l'affichage suivant :

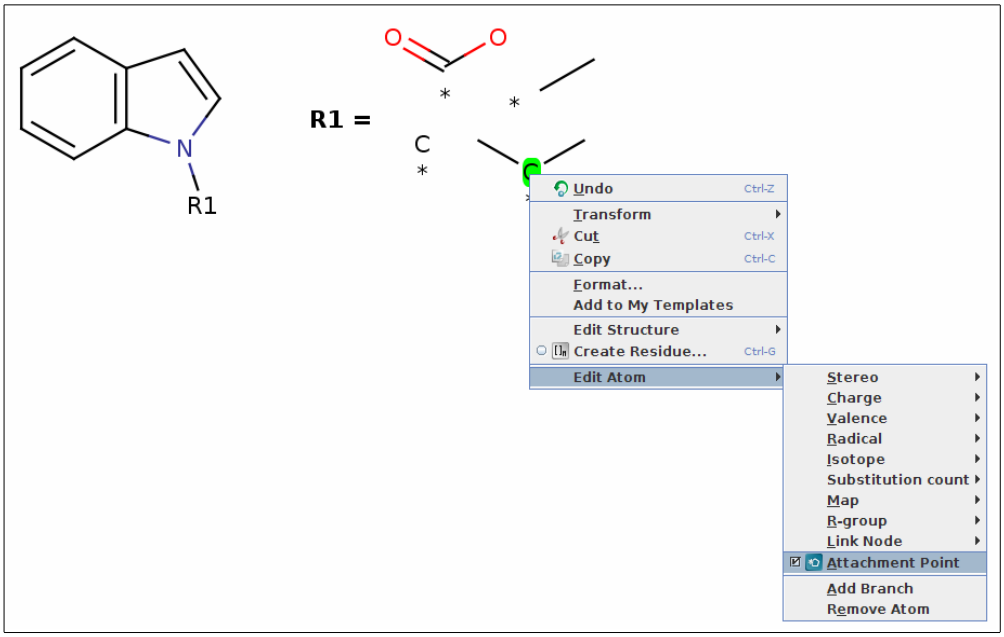


C - Définition des points d'attachement

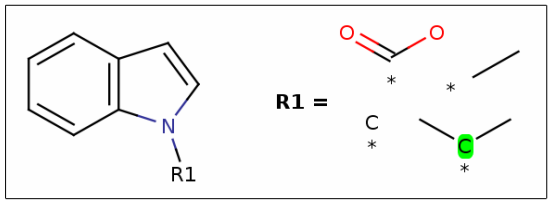
Il convient maintenant de désigner l'atome de chaque fragment qui sera lié à l'atome d'azote pour que la requête soit correctement comprise dans l'interface de requête.

On sélectionne un par un chaque atome et on active la commande **Edit Atom/Attachment Point** de la Figure 28.

Figure 28



L'affichage devient :



Astuce : si on a appliqué l'astuce vue à l'étape B, le curseur actif indique toujours R1 ; il suffit alors de cliquer sur un atome pour que celui-ci devienne point d'attachement.

Si on pose cette question en cochant **As drawn** dans l'interface de requête ET en laissant les autres options non cochées, on obtient 42 substances.

C-VIII-2-b- Deux points d'attachement

Cette procédure était, jusqu'en septembre 2011, nécessaire pour définir des chaînes ou des cycles de taille variable. Actuellement ce problème est facilement résolu (cf. C-VII-2).

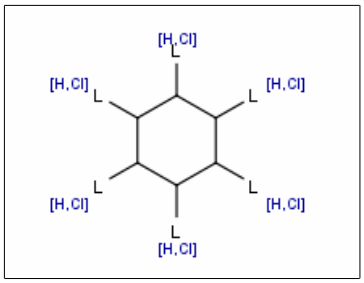
Il est toutefois imaginable de construire une requête plus complexe avec deux, voire plus, points d'attachement sur un ou plusieurs fragments.

C-VIII-3- Groupements variables répétés n fois

On s'intéresse ici à une sous-structure qui comporte un groupement variable présent n fois. Il y a deux procédures possibles :

C-VIII-3-a- La plus directe

Par ex., si l'on cherche l'ensemble des cyclohexanes dichlorés sans plus de précision sur la position relative des 2 atomes de Chlore sur le cycle, on dessinera un cyclohexane avec 6 liaisons reliées à une liste d'atomes (cf. C-III-2) composée des atomes H et Cl.

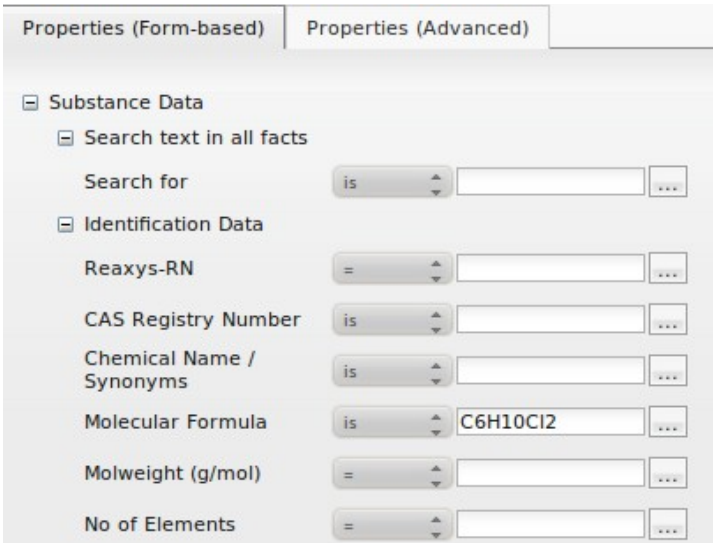


Pour préciser que l'atome de Chlore n'est présent que deux fois sur la sous-structure, on transfère cette requête en mode **Query** et on rajoute cette contrainte par l'intermédiaire du formulaire **Properties (Form-Based)** montré ci-dessous. On cherche la propriété **Molecular Formula** en développant l'arborescence des champs de la base (cf. B-I et Fig. 29) et on saisit la valeur :

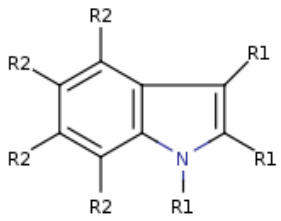
```
C6H10Cl12*
```

Cette requête, avec les options **As drawn** et élimination des autres options, conduit à 14 substances répondant à la question.

Figure 29



C-VIII-3-b- La plus largement applicable



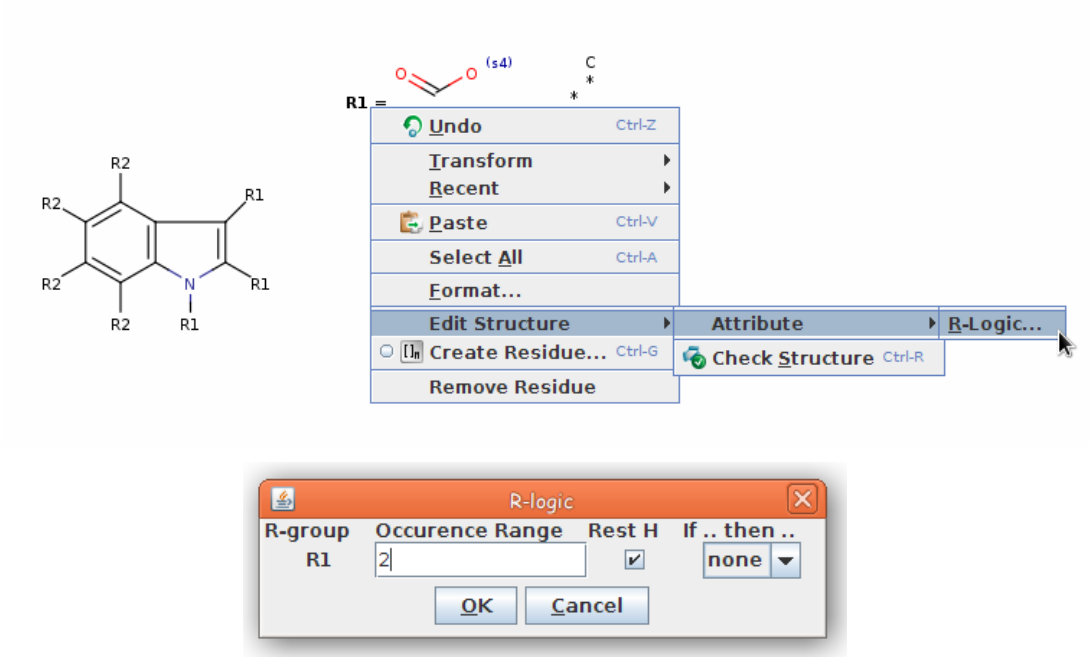
On va s'intéresser à des dérivés de l'indole comportant deux groupements génériques R_1 et R_2 répétés plusieurs fois et comportant eux-mêmes plusieurs sous-structures (cf. Figure ci-contre). on devra appliquer la procédure ci -dessous.

Le groupement R_1 sera identique à celui de l'exemple abordé en C-VIII-2-a et positionné le cycle azoté de l'indole et sa fréquence fixée à 2. Le groupement R_2 sera une liste d'halogènes (Cl, Br, I) et sa fréquence fixé à 2 également.

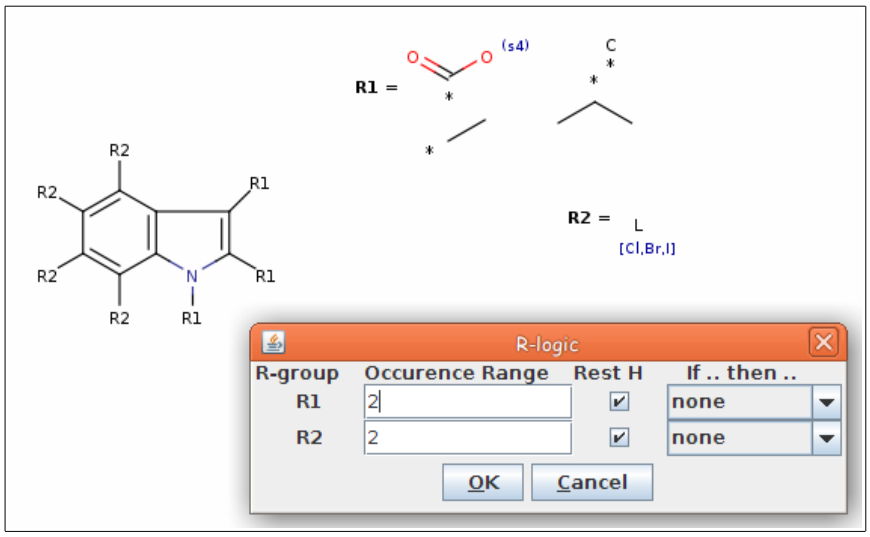
Une fois les étapes A, B et C (expliquées en C-VIII-2-a) appliquées au groupe R1, on définit la fréquence de Markush en effectuant un clic-droit sur le symbole **R1** (Fig. 30) puis en activant la commande **Edit Structure/Attribute/RLogic**. on saisit la valeur '2' tel qu'indiqué sur la boîte de dialogue **R-logic**, en cochant la case **Rest H** pour signifier qu'en l'absence du groupement R1, il n'y aura pas d'autre substituant que l'Hydrogène.

L'affichage du dessin reste inchangé et on ne voit pas graphiquement de modification résultant de l'attribution de la fréquence.

Figure 30

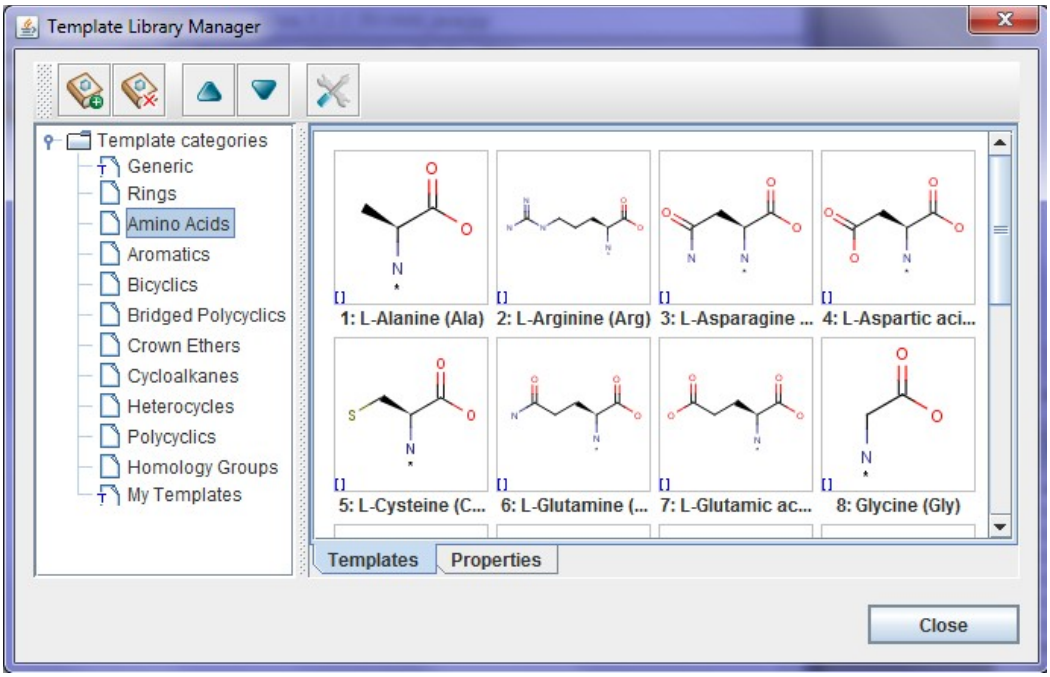


On effectue les mêmes opérations pour définir en R_2 la liste d'halogènes (Cl, Br, I) en lui attribuant une fréquence de 2 ; pas besoin de lui attribuer un point d'attachement puisqu'il ne s'agit que d'un seul atome générique. La Figure ci-dessous illustre le schéma dans sa phase quasi-finale où l'on définit la fréquence de R_2 à 2.

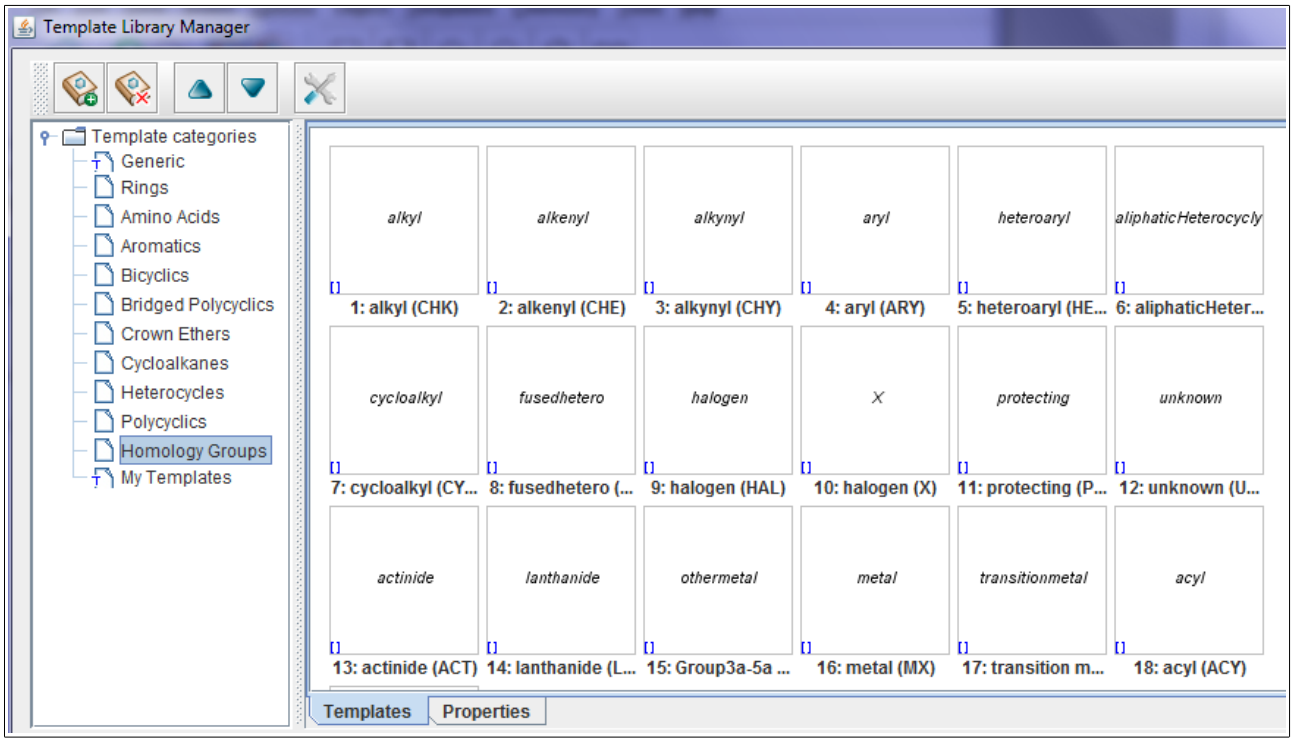


C-IX- Templates

Outre les outils classiques de dessin (atomes du tableau périodique, atomes usuels, liaisons avec ou sans stéréochimie), MarvinSketch propose deux types de **Templates**, accessibles par le menu du même nom : les groupements génériques **Reaxys Generics** – déjà vue au paragraphe C-VIII-1 – et une bibliothèque de structures pré-dessinées, la **Template Library**,



Dans cette liste de catégories, l'une d'entre elles n'est pas courante dans les éditeurs de structure usuels; il s'agit de la catégorie **Homology Groups**.



D – INTERROGATION PAR STRUCTURE

Remarques préliminaires :

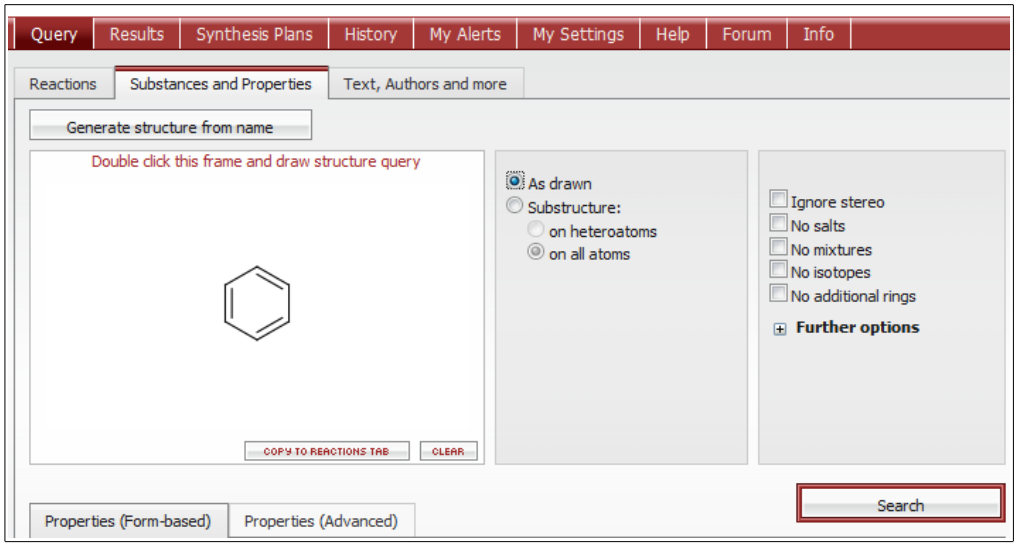
- Parmi les différentes façons de chercher une substance : nom trivial, nom IUPAC, formule, recherche par structure... c'est cette dernière méthode qui reste la plus sûre dans de nombreux cas.
- Les requêtes et l'historique associé sont conservés par défaut sur le serveur jusqu'à déconnexion de la session. On peut toutefois les sauvegarder sur son poste de travail en ouvrant le mode **History** (du bandeau principal) qui conserve une copie de la formulation de la question ainsi que les listes de réponses (cf. G-III). Le résultat d'une recherche dépend largement de la formulation de la question : un petit oubli peut avoir des conséquences importantes.

D-I- Recherche par structure exacte

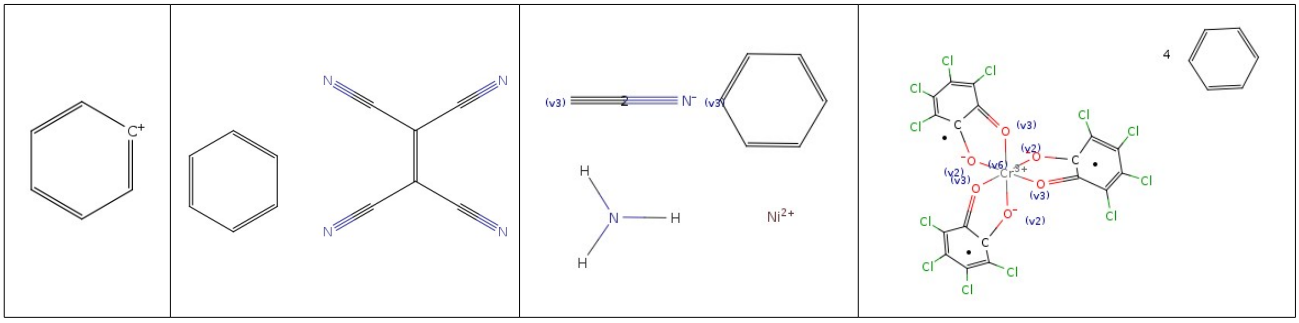
Concernant la recherche d'un composé, la requête par structure sera la plus sûre, la recherche par nom n'étant pas univoque. Il suffit de dessiner la molécule dans l'éditeur de structures et de la transférer dans l'interface de requête. Évidemment il faudra tenir compte des options présentes à droite de l'interface.

Par ex. si on interroge avec la molécule de benzène et les options telles que sur la Fig. 31,

Figure 31



on obtient 8005 substances, malgré le choix **As drawn** dans la partie gauche des options, car les autres options de requête autorisent la présence de sels, mélanges, isotopes, cycles supplémentaires. Par exemple on retrouve des molécules telles que :



Si l'on veut ne retrouver que la molécule de benzène, il faudra cocher toutes les options **Ignore stereo, No salts, …** y compris les options **No charges** et **No radicals** qui se trouvent dans **Further options** (Fig. 31).

Dans la majorité des cas, sur une longue liste de réponses obtenue sans tenir compte des options, il suffit de trier la liste par nombre de références décroissant pour que la première soit la substance recherchée.

D-II- Recherche par sous-structure

On dispose pour dessiner des outils qui ont été décrits en section C.

Il faudra, pour arriver rapidement à des réponses pertinentes bien gérer les degrés de substitution des différents atomes. Nous distinguerons dans les paragraphes suivants deux types de substances :

- soit des substances avec des substitutions variables,
- soit des substances contenant plusieurs sous-structures distinctes.

D-II-1- Sous-structures avec substitutions variables

Option ***As drawn***

On peut poser des requêtes pour lesquelles les substitutions ne sont permises que sur certains atomes; les outils de dessin appropriés sont décrits au paragraphe C-III-1.

Si pour l'exemple de l'aspirine ci-contre, seuls les atomes d'Oxygene et de Carbone doivent être substituables, alors on peut poser la question correspondante dans l'interface de requête en choisissant l'option **As drawn** et les exclusions de la partie droite de la Fig. 32. On obtient dans ce cas 1263 dérivés de l'aspirine valides par rapport à la requête.

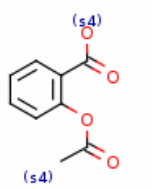
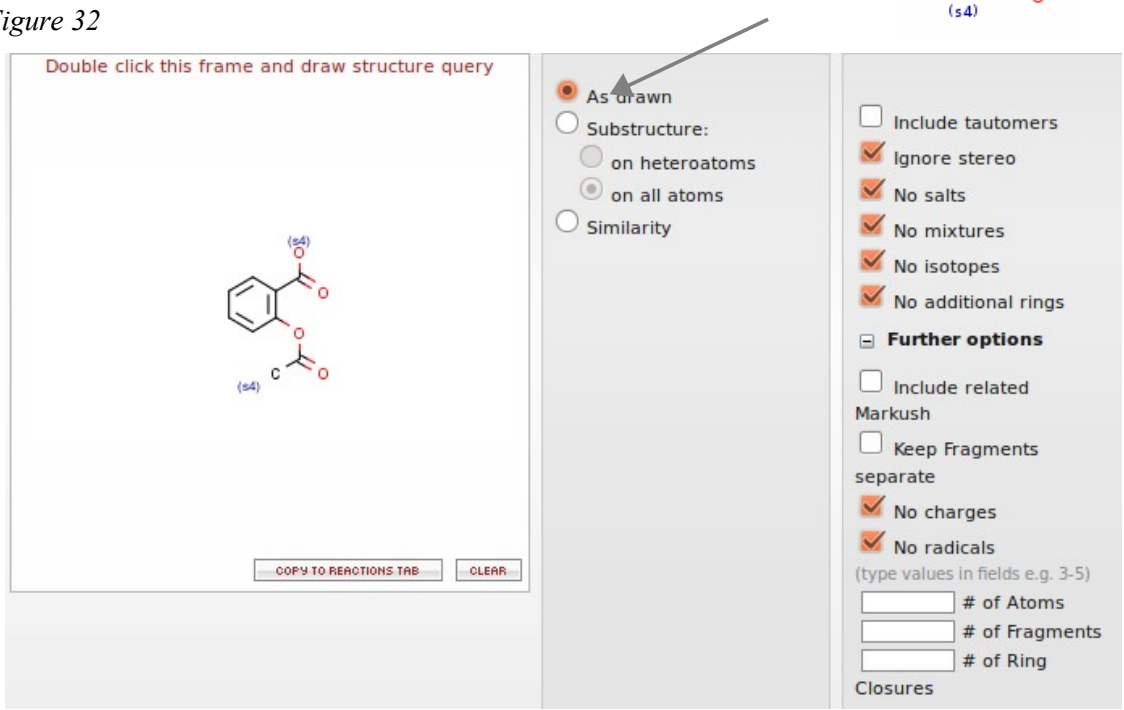
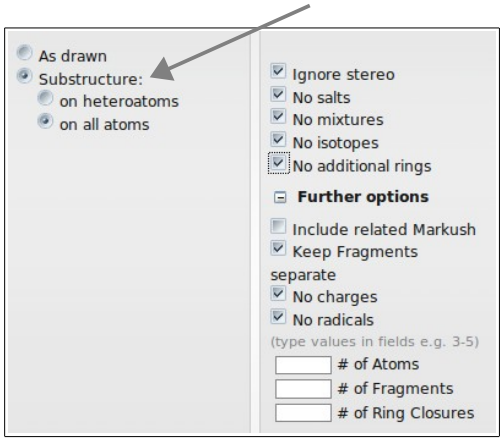


Figure 32

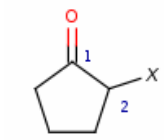
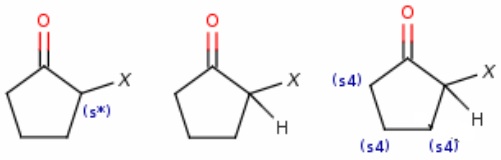
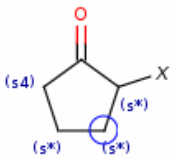


Option ***Substructure***

Deux possibilités : recherche sous-structurale sur les hétéroatomes seulement ou sur tous les atomes.



On peut également préciser le degré de substitution d'atomes spécifiques et interroger avec cette option **Substructure on all atoms**. Supposons que l'on cherche des dérivés de la cyclopentanone possédant un halogène en position α. Dans l'interface de requête, nous rechercherons par sous-structure (présence de l'halogène oblige) avec les options de requête telles que précisées ci-contre.

1er exemple : toutes les positions de la cyclopentanone sont substituables. Requête par **Substructure**.		526 réponses
2ème exemple : la position 2 n'est pas substituable; les 3 autres positions le sont. Il y a au moins 3 manières équivalentes de formuler cette requête. Requête par **Substructure**.		253 réponses
3ème exemple : les positions 2, 3 et 4 sont bloquées; la position 5 est substituable. Requête avec l'option **As drawn**		53 réponses

Si l'on pose la requête comme dans le 1er exemple mais en cochant le bouton **As drawn** (mêmes options) dans l'interface de requête, alors on n'obtient plus que 4 réponses correspondant aux 4 halogènes courants.

Autres options de recherche

Il s'agit des options situées dans la partie droite de la Fig. 32 par exemple :

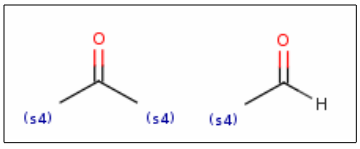
- **Ignore stereo** : s'applique aux Carbones sp^2 ou sp^3. Les notions de stéréochimie absolue ou relative ne sont pas prises en compte lors des requêtes (cf. C-VI).
- **No salts, mixtures, isotopes** : pas de commentaires
- **No additional rings** : intervient sur la présence ou non de cycles supplémentaires directement attachés à la molécule-requête. Ne pas confondre avec l'option **Number of Ring closures** ci-dessous.
- **Include Related Markush** : une structure de Markush est une description d'une classe de composés par une notation générique. Elle est souvent utilisée dans les revendications des brevets et pour des librairies combinatoires. Les groupements variables (cf. C-VIII), listes d'atomes lists, avec ou sans répétition sont fréquemment utilisées dans ces représentations de Markush.
- **Keep fragments separate** : utile pour une recherche simultanée de fragments distincts (cf. D-II-2)
- **No charges, radicals** : pas de commentaires
- **Number of atoms, fragments** : pas de commentaires
- **Number of Ring closures** : utile pour rechercher des cycles d'une taille spécifique. Le **Number of Ring Closures** reflète le nombre total de cycles de la sous-structure.
- **Recherche automatique** : c'est une option qui permet en cas de non-réponse à une requête simple d'étendre automatiquement la recherche avec substituants possibles sur les hétéroatomes dans un premier temps, sur tous les atomes dans un deuxième temps.

Quelle peut être l'utilité de cette option ? En cas de recherche précise, aucune. Par contre elle peut être utile pour ne pas passer à côté de bonnes réponses si l'on n'est pas sûr de la formulation de sa requête. L'inconvénient de cette option, lorsqu'elle est active, est qu'elle ralentit le temps de réponse à une requête donnée. On peut choisir de désactiver ou non cette option par le menu **My Settings** de l'interface principale à la rubrique **Substance search options**.

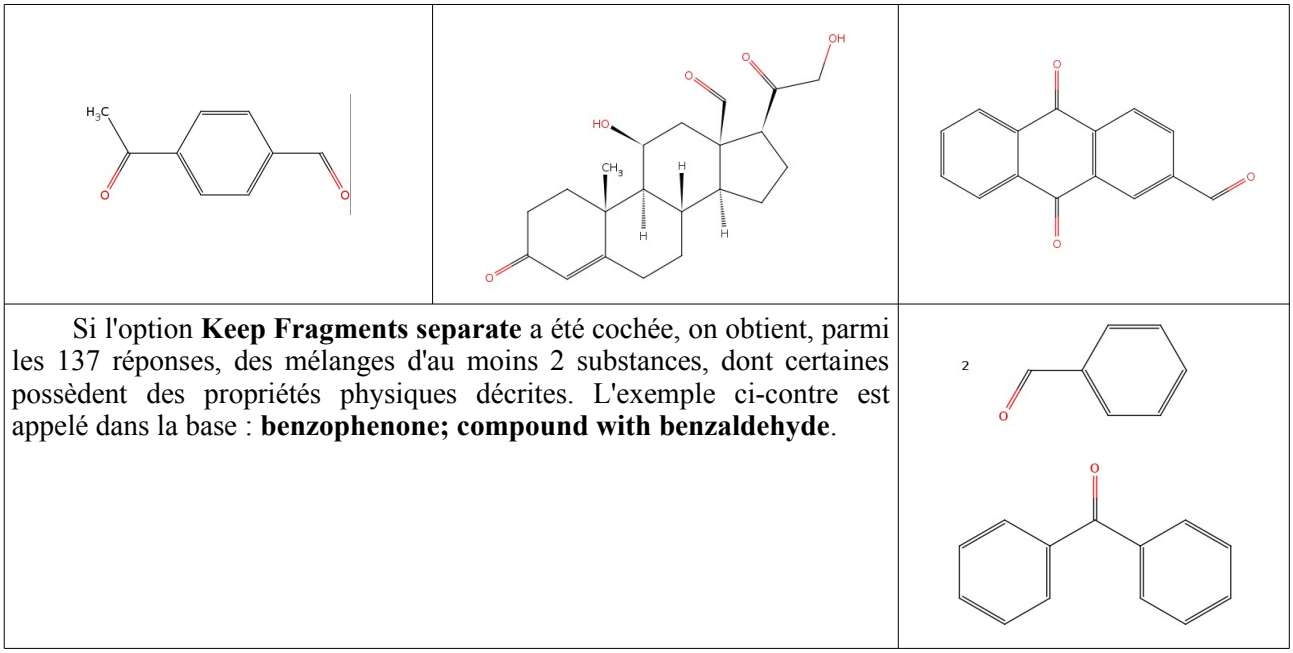
Remarque : cette option n'est pas visible dans l'interface de requête.

D-II-2- Substances contenant plusieurs sous-structures

On peut rechercher des substances contenant des fragments distincts : une cétone ET un aldéhyde par exemple.



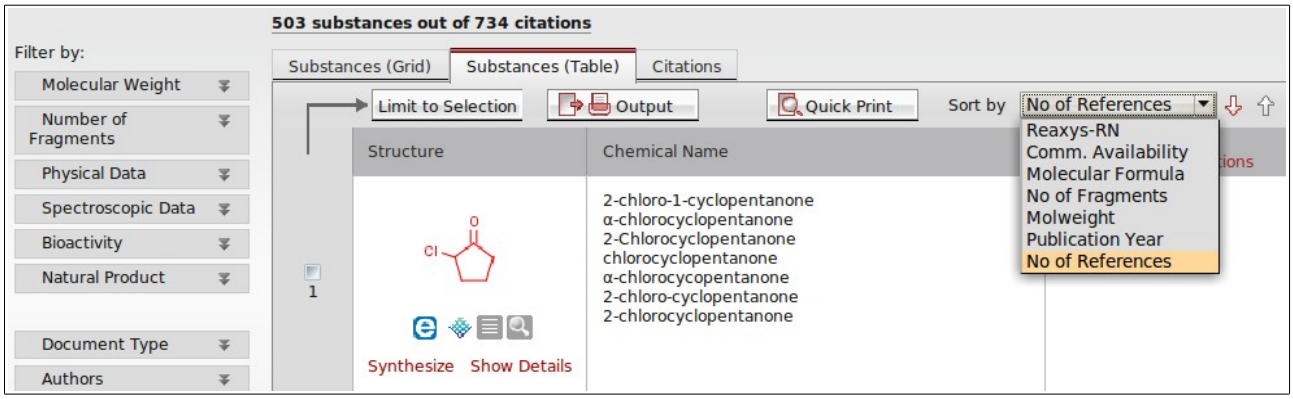
C'est une requête très large qui fournit – en autorisant les cycles supplémentaires dans les options – 19 475 réponses. On constate que les deux fonctions cétone et aldéhyde appartiennent à la même molécule si l'option **Keep Fragments separate** est restée inactive.



Si l'option **Keep Fragments separate** a été cochée, on obtient, parmi les 137 réponses, des mélanges d'au moins 2 substances, dont certaines possèdent des propriétés physiques décrites. L'exemple ci-contre est appelé dans la base : **benzophenone; compound with benzaldehyde**.

D-III- Tri des réponses

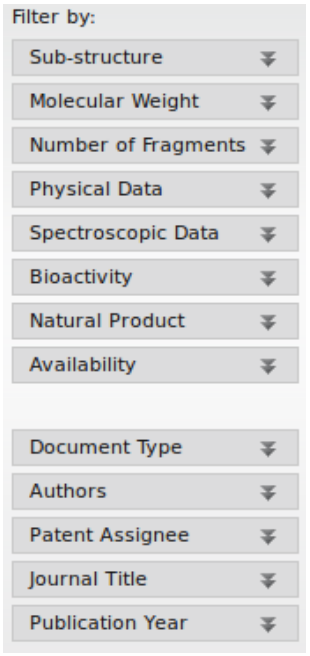
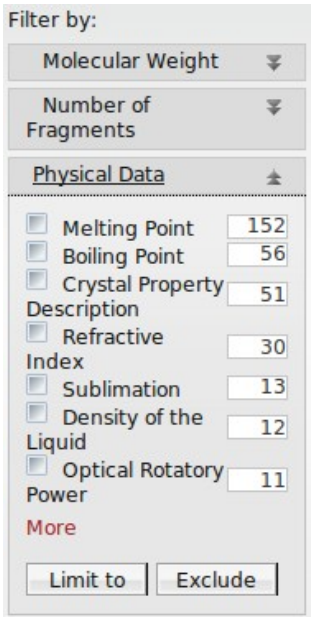
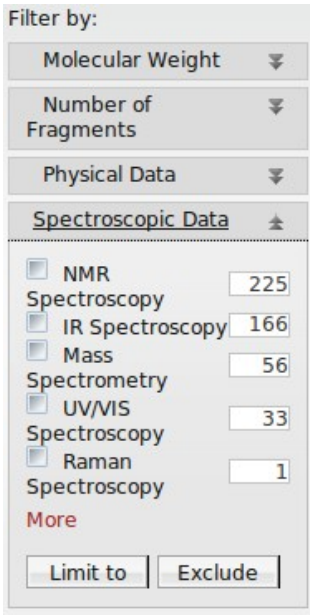
Premier type de tri : comme indiqué sur la copie d'écran ci-dessous, les options de tri (Reaxys-RN, …., Nombre de références) permettent seulement un affichage différent. L'option par défaut est le nombre de références qui donne déjà un premier tri pertinent.



Deuxième type de tri : les filtres qui eux, par contre, vont restreindre le nombre de réponses sur des critères généraux montrés sur la colonne Filtres généraux de la Fig. 33. Deux filtres sont explicités sur cette figure : le filtre par donnée physique et par donnée spectrale.

On peut sur chaque filtre limiter le nombre de réponses correspondant au critère choisi ou exclure ce même nombre de réponses.

Figure 33

Filtres généraux	par donnée physique	par donnée spectrale
		

E – INTERROGATION PAR SCHÉMA RÉACTIONNEL

Les requêtes sur les schémas réactionnels peuvent être faites par structure ou par sous structure.

E-I- Dessin des structures

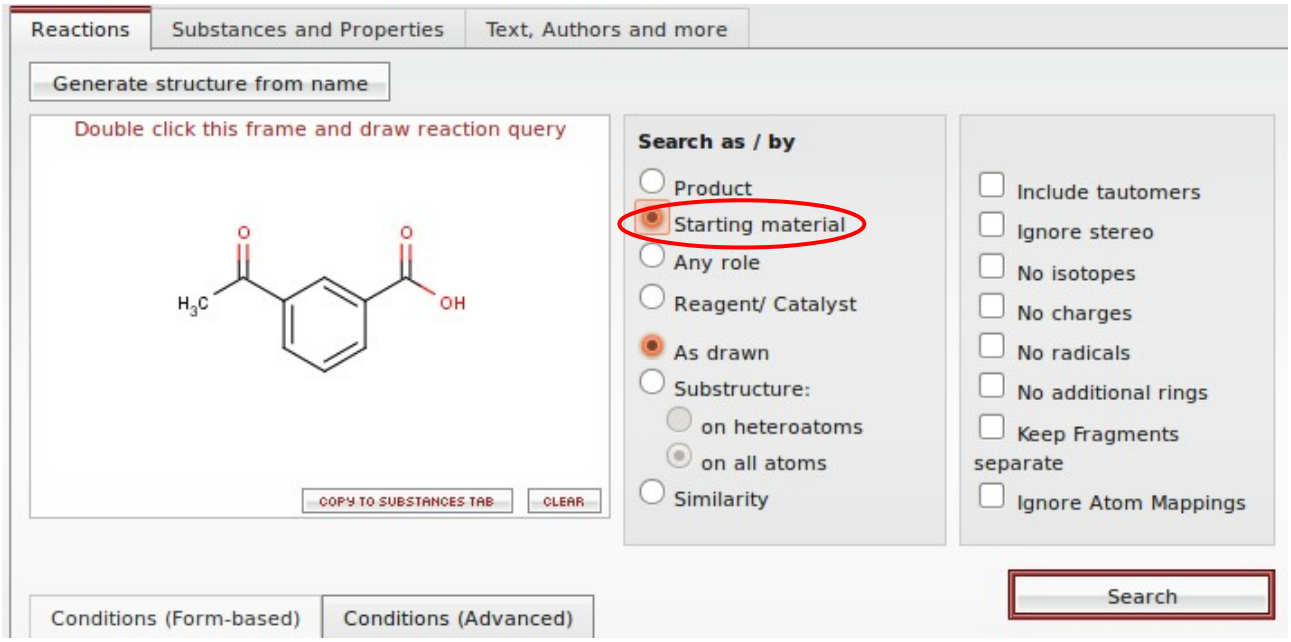
Le dessin des réactants et/ou produits de la réaction est effectué dans le **Structure Editor** comme indiqué au chapitre E. Les réactifs, catalyseurs, solvants ne sont pas indiqués dans le schéma réactionnel lui-même.

E-II- Schéma réactionnel

E-II-1- Schéma réactionnel incomplet (half reaction)

Pour dessiner une requête ne précisant que le produit de départ (**starting material**) ou le produit d'arrivée (**product**) en structure exacte ou sous-structure, il n'est pas indispensable de dessiner un schéma réactionnel complet dans l'éditeur de structure. Il suffit de préciser la structure et de simplement cocher l'option correspondante dans les options de la requête comme sur la Fig. 34. Cette requête fournit 71 réactions qui reflètent la réactivité de certains atomes ou fonctions de la molécule de départ.

Figure 34



La requête ci-dessus s'effectuera dans les schémas réactionnels complets et les demi-réactions, les **Half reactions** indexées dans la base (cf. A-II-2) Evidemment on ne pourra pas utiliser des indications de mapping (cf. E-III-1) compte-tenu de la nature de la requête, restreinte à un composant

E-II-2- Schéma réactionnel complet

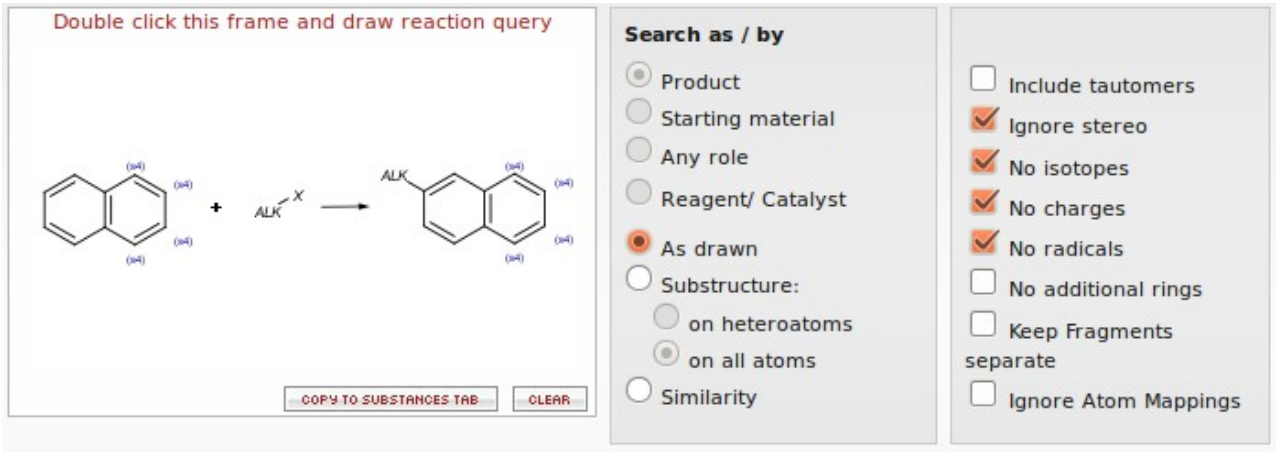
La procédure et les différentes options vont être indiquées en se servant d'un exemple.

Soient à rechercher les réactions d'alkylation de type Friedel et Crafts, impliquant l'addition d'un halogénure d'alkyle sur une sous-structure naphtalénique (Fig. 35). L'un des cycles peut être substitué par n'importe quel groupement, y compris des cycles, l'alkylation se faisant sur l'autre cycle en position 2.

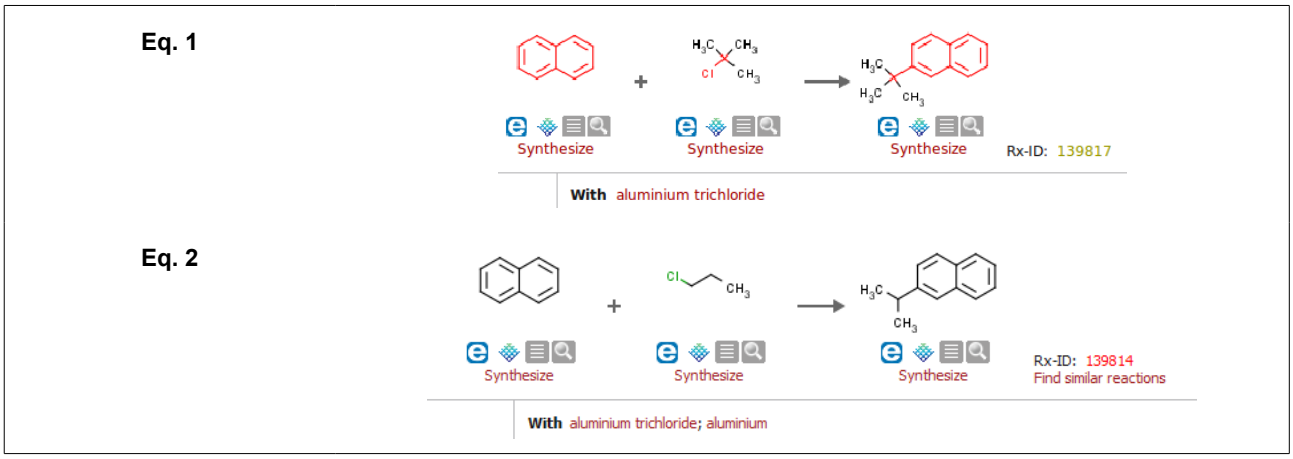
Les réactant(s) et produit(s) sont dessinés avec les spécifications indiquées; on rappelle que les groupes alkyle et halogénure sont accessibles par les **Templates Reaxys** (cf. C-II-1 et C-IX) et que le degré de substitution des atomes doit être fixé comme vu au paragraphe C-III-1.

Remarque : le signe + se met automatiquement entre deux produits de départ ou d'arrivée, une fois que la flèche de réaction est dessinée.

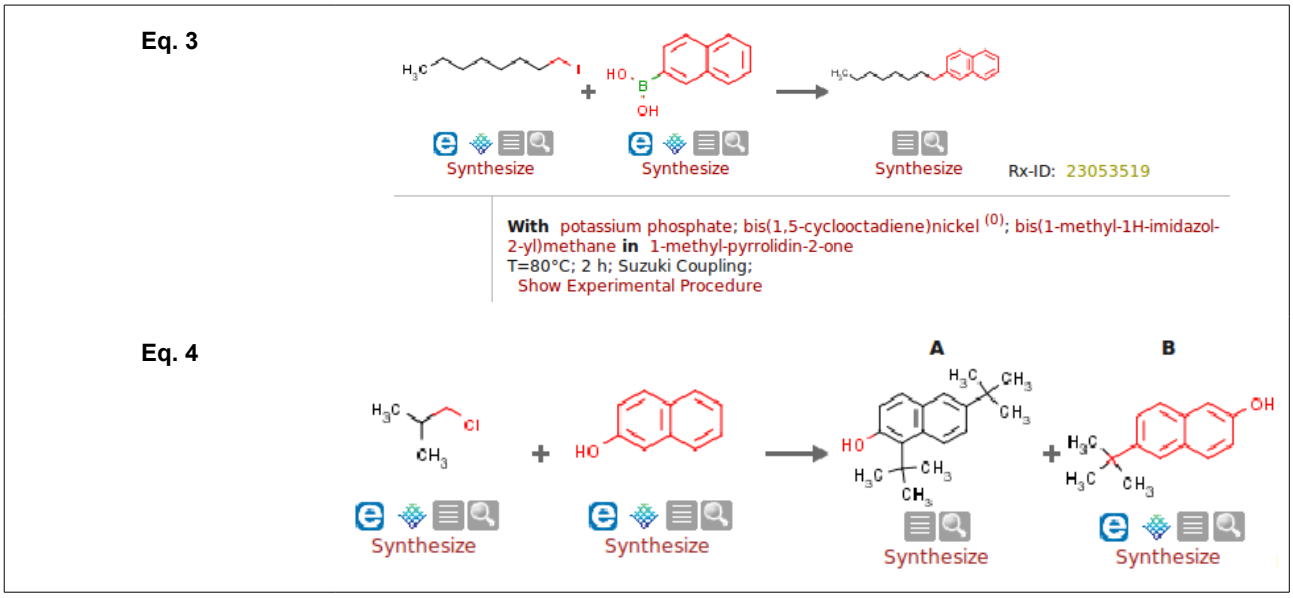
Figure 35 : alkylation de type Friedel et Crafts avec les options correspondantes



La question formulée avec les options ci-dessus conduit à 77 réponses, dont certaines sont pertinentes (Eq. 1 et 2), mais aussi à des réponses qui ne correspondent pas à ce que l'on aurait souhaité.



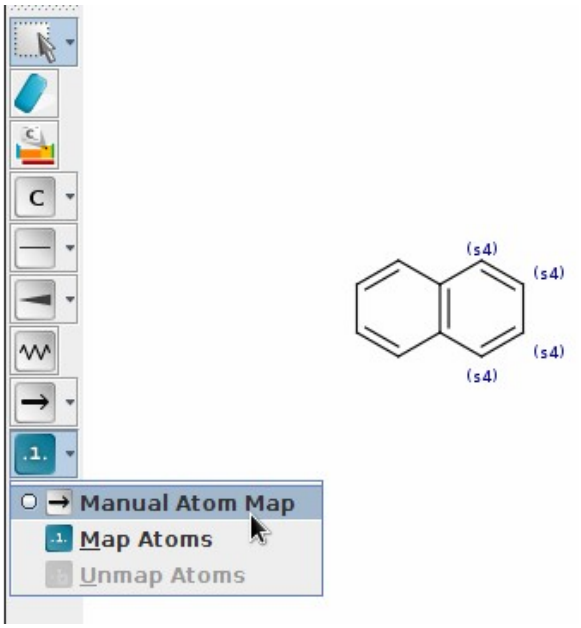
La réaction de l'équation 3 n'est pas une réaction de Friedel et Crafts mais un couplage de type Suzuki. Enfin l'équation 4 montre une réaction de diaddition.



On peut essayer d'éliminer ces réactions en réalisant le **mapping** de la réaction-requête (cf. E-III-1).

E-III- Spécifications des requêtes par schéma réactionnel

E-III-1- Mapping

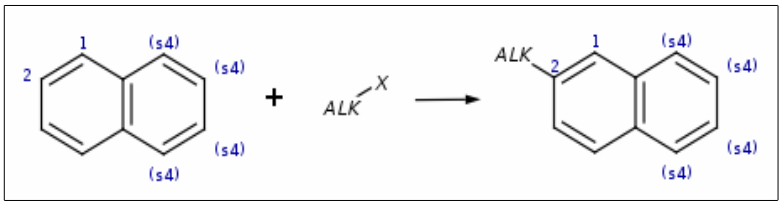


C'est par définition l'indication de la correspondance, ou appariement, entre les atomes du(des) réactant(s) et du(des) produit(s). On peut mapper les atomes et aussi, le cas échéant, les groupements G variables (cf. E-IX).

La précision du mapping, dans l'exemple du paragraphe E-II-2, va permettre d'éliminer des réactions non pertinentes obtenues dans la première requête. Pour spécifier le mapping, on utilise l'outil montré sur la copie d'écran ci-contre.

Dans cet exemple, on utilisera plutôt l'outil **Manual Atom Map** qui permet un meilleur contrôle de l'appariement. Concrètement on fait un "glisser-tirer" entre les atomes devant être appariés.

Avec la requête telle que dessinée ci-dessous, seules 26 réactions sont retrouvées si l'on effectue un mapping minimal, à partir des atomes C1 et C2, on a éliminé le couplage de Suzuki (Eq. 3); si l'on précise le mapping sur le groupement alkyle, on n'obtient plus que 20 réponses; les réactions de diaddition (Eq. 4) ont aussi disparu.



Remarque : *dans Beilstein, de nombreuses réactions (environ 30 %) sont indexées sans mapping, d'autres le sont avec un mapping chimiquement faux. Une requête avec mapping ne permettra pas d'accéder aux réactions non mappées ou mappées avec un mapping différent mais elle sera plus précise et a priori plus pertinente.*

Remarque *: il n'est pas utile de mapper tous les atomes.*

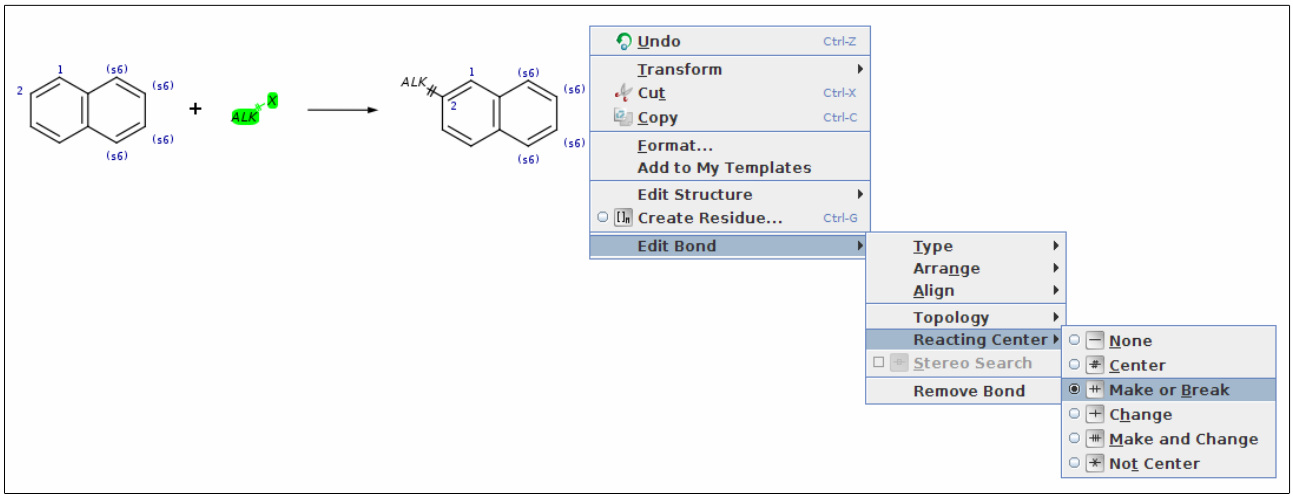
Le mapping peut être effacé par la commande **Unmap Atom** de la Fig. 13.

E-III-2- Centres réactionnels

Outre le mapping sur les atomes, il peut être utile de préciser quelles liaisons sont cassées ou formées durant la réaction, ou autres modifications du statut des liaisons. Dans l'exemple en cours, on sélectionne la liaison comme indiqué sur la Fig. 36 pour préciser que la liaison ALK–X est cassée dans le produit de départ et que la liaison ALK–C2 est formée dans le produit d'arrivée.

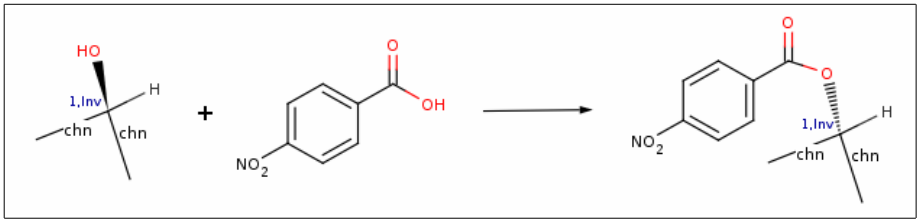
On retrouve les 20 réactions déjà obtenues précédemment.

Figure 36



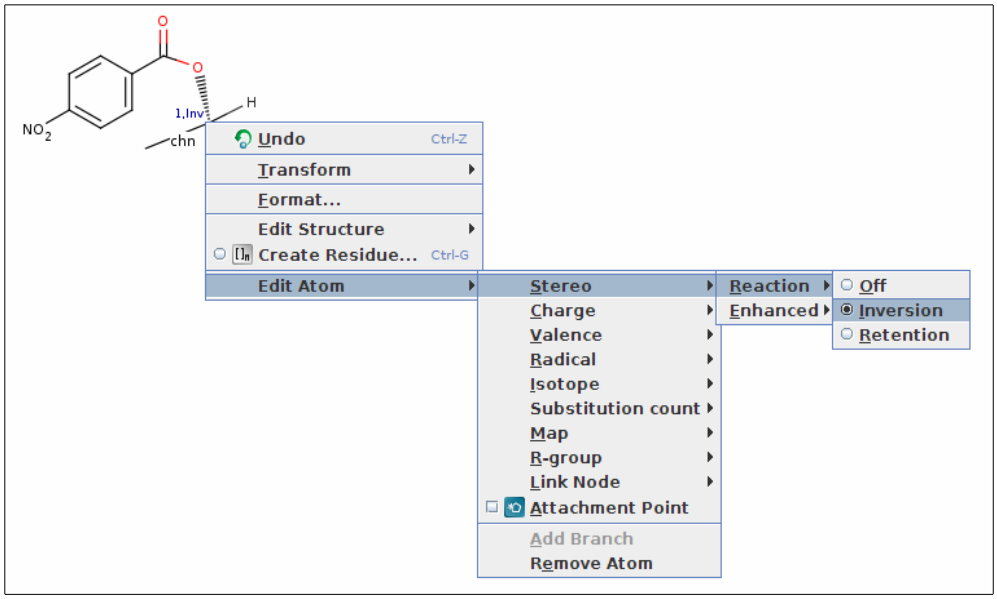
E-III-3- Cours stéréochimique des réactions

En mode réaction, le crayon étant activé, on clique sur un atome et on peut alors préciser le cours stéréochimique (inversion/rétention) au niveau d'un centre réactionnel avec les options **Inversion** et **Retention.** Ci-dessous est donné le schéma d'une réaction dite de Mitsunobu qui permet de réaliser une inversion de configuration sur le Carbone porteur de la fonction hydroxyle.



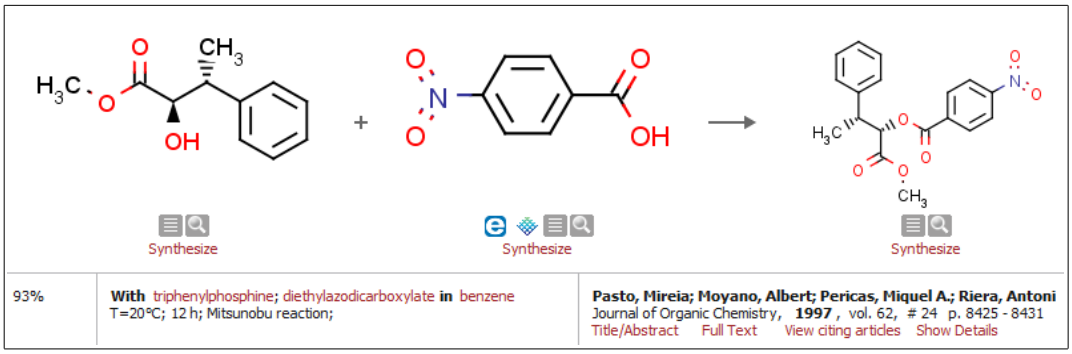
En pratique on utilise les menus contextuels à partie d'un clic-droit sur l'atome concerné pour développer les menus comme indiqué sur la Fig. 37.

Figure 37



Si l'on interroge Reaxys avec les options de la Fig. 31, on trouve 7 réactions qui répondent à la question; une des réponses est donnée sur la Fig. 38 ci-dessous.

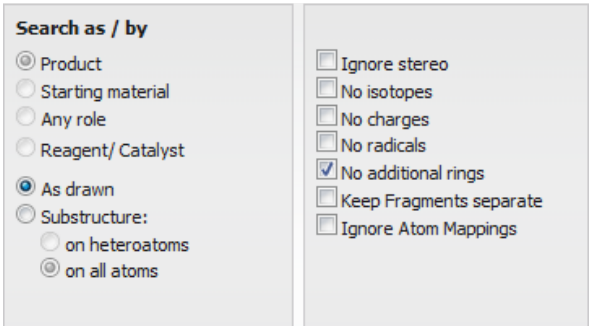
Figure 38


Synthesize
Synthesize
Synthesize
93%
With triphenylphosphine; diethylazodicarboxylate in benzene
T=20°C; 12 h; Mitsunobu reaction;
Pasto, Mireia; Moyano, Albert; Pericas, Miquel A.; Riera, Antoni
Journal of Organic Chemistry, 1997, vol. 62, # 24 p. 8425 - 8431
Title/Abstract Full Text View citing articles Show Details

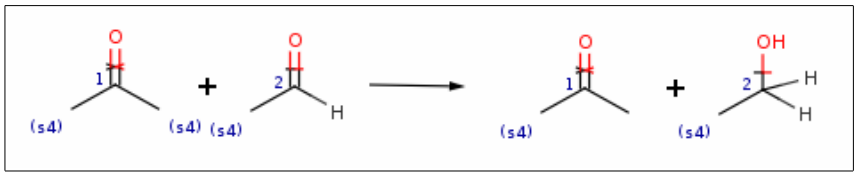
E-IV- Réactions sélectives

Outre le type de requête exemplifié ci-dessus, il existe diverses possibilités de formulation des requêtes par schéma réactionnel. Une de ces possibilités concerne les réactions sélectives qui sont très souvent requises lors de la synthèse d'un composé multi-fonctionnel.

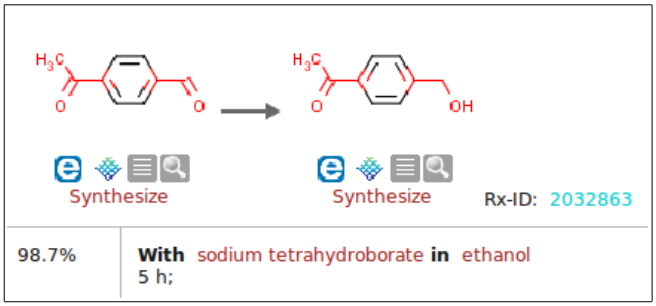
On cherche, par ex., les réactions de réduction chimiosélective d'un aldéhyde en présence d'une fonction cétone contenue dans la même molécule. On retiendra l'option **No additional rings** pour réduire le nombre de réponses comme indiqué ci-dessous.



La requête peut se dessiner, comme montré ci-dessous, avec les spécifications de mapping minimal et les changements de statut des liaisons (cf. E-III-2); le carbonyle de la cétone reste inchangé alors que celui de l'aldéhyde est modifié.



On obtient 35 réactions qui semblent pertinentes du point de vue de la chimie et qui sont majoritairement des réactions intramoléculaires si l'on ne considère que les produits de départ. On observe quelques réactions conduisant à 2 produits : le produit attendu et le produit de sur-réduction de la cétone.



En réunissant les fragments cétone et aldéhyde par un groupement générique G (voir **Templates Reaxys**, cf. C-IX), on n'obtient que des réactions intramoléculaires et des réactions où les fonctions aldéhyde et cétone sont séparées par un cycle ou un hétéroatome.

E-V- Réactions multi-étapes

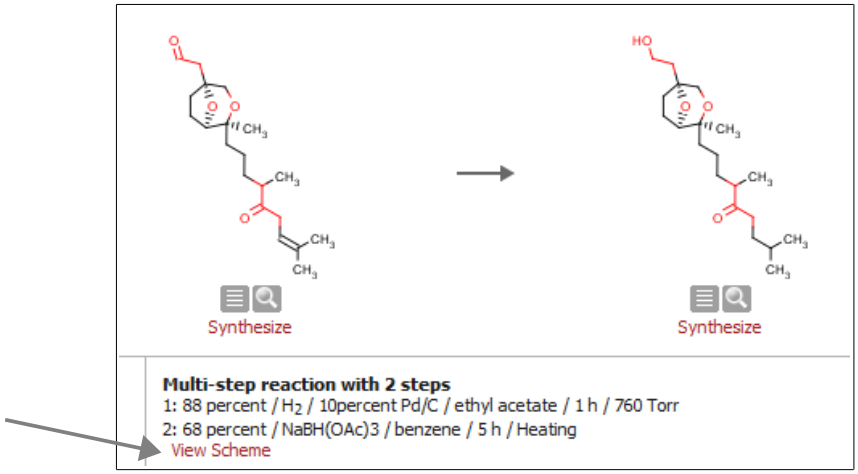
Comme le CAS, Reaxys indexe aussi les réactions multi-étapes faisant partie d'une séquence en considérant toutes les combinaisons possibles; par exemple, pour la séquence A→B→C→D se trouvant dans un même document, non seulement chaque réaction A→B ; B→C ; C→D est susceptible d'être indexée, mais aussi les combinaisons : A→C ; A→D et B→C. Ceci peut présenter des avantages concernant l'indexation de méthodes de synthèse mais provoque une « augmentation artificielle » du nombre de réactions indexées.

Ces réactions sont reconnaissables dans l'affichage des résultats; la Fig.39 montre une réaction multi-étapes comportant deux réactions mono-étapes.

Remarque : On peut obtenir une description de chaque schéma réactionnel en activant le lien ***View Scheme*** *ce qui permet de visualiser chaque réaction individuelle en mode* ***Synthesis Plans*** *(cf. E-VII).*

L'intérêt de ces réactions multi-étapes réside dans la possibilité de retrouver des méthodes de synthèse, sachant que les 11,6 M de réactions impliquées dans ces séquences recouvrent environ 6% du total des citations de la base.

Figure 39


Synthesize
Synthesize
Multi-step reaction with 2 steps
1: 88 percent / H2 / 10percent Pd/C / ethyl acetate / 1 h / 760 Torr
2: 68 percent / NaBH(OAc)3 / benzene / 5 h / Heating
View Scheme

Ces réactions multi-étapes permettent d'accéder éventuellement à des séquences réactionnelles plus longues. La séquence la plus longue contient 54 réactions individuelles !

Pour finir, on peut fixer dans sa recherche un nombre d'étapes ou une plage de valeurs correspondant à des réactions multi-étapes. Le champ correspondant est RXD.STP ou **Number of reaction steps** (cf. section B-I-1 pour la recherche par champs). On peut aussi les visualiser en effectuant un tri dans les réponses (cf. E-VI).

E-VI- Tri des réponses

Il y a pour les réactions deux modes différents pour trier une liste de réactions :

- le premier (Fig. 40), en haut de l'interface, avec en particulier l'option par défaut **Reaxys-Ranking**. Bien que l'algorithme de tri correspondant ne soit pas public, il semble donner de bons résultats en partie calculés sur le rendement. Ces options de tri ne modifient que l'affichage des réponses
- le deuxième, sur le côté gauche de l'interface, propose plusieurs options de tri, qui modifient le nombre de réponses :
  - par sous-structure (Fig. 40a); ceci permet de restreindre la liste côté produits de départ ou d'arrivée.
  - sur le rendement
  - par type d'enregistrement (Fig. 40b)
  - sur la nature du réactif ou du catalyseur (Fig. 40c), du solvant
  - par type de réaction (Fig. 40c)
  - par nombre d'étapes.

Figure 40

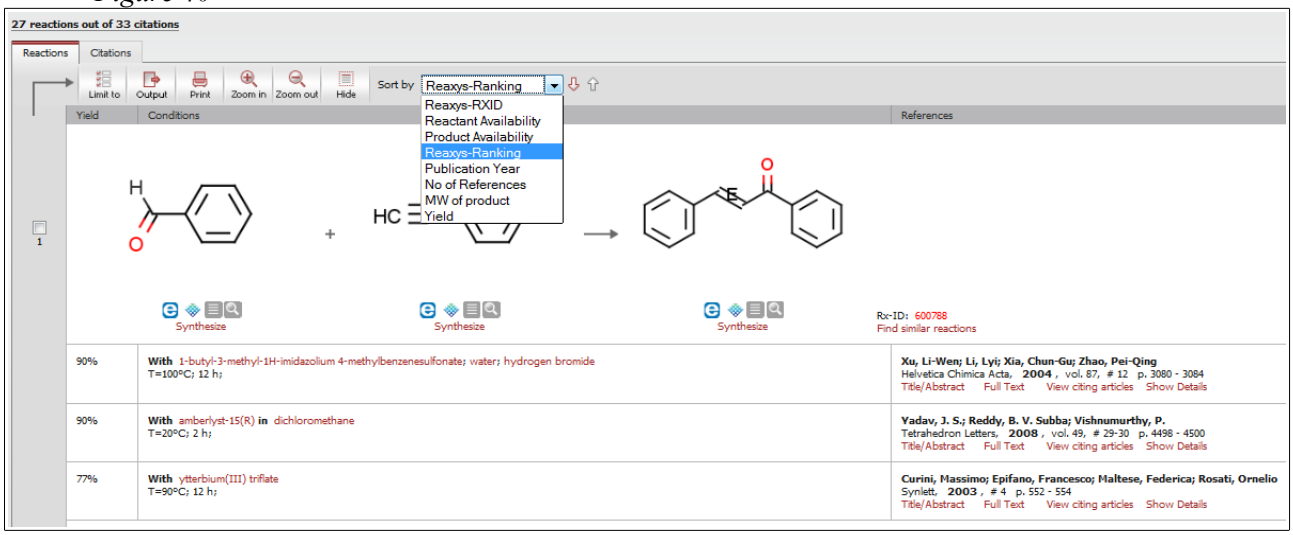
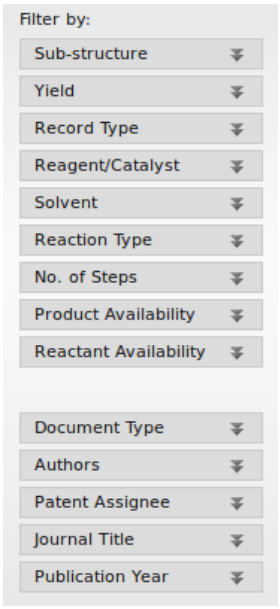
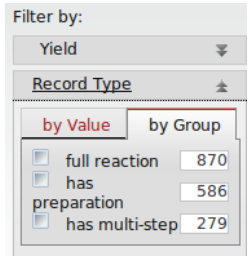
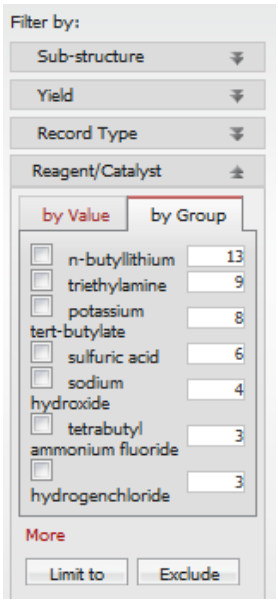
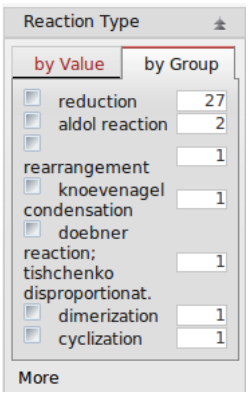
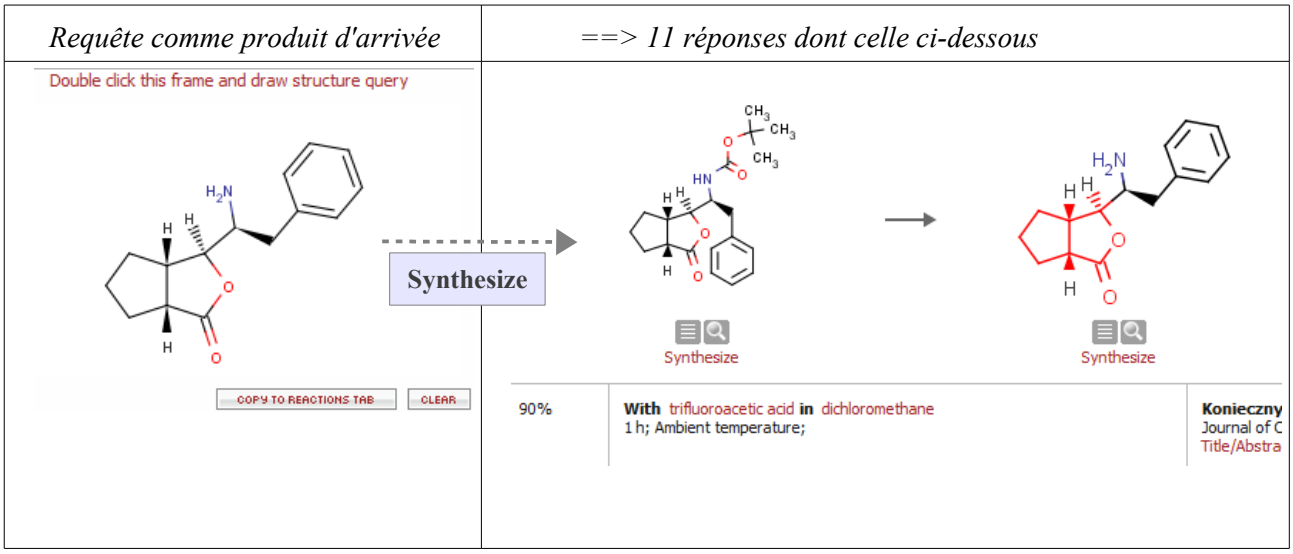


Figure 40a	*Figure 40b*	*Figure 40c*	*Figure 40d*
			

E-VII- Plans de synthèse – Aide à la rétrosynthèse

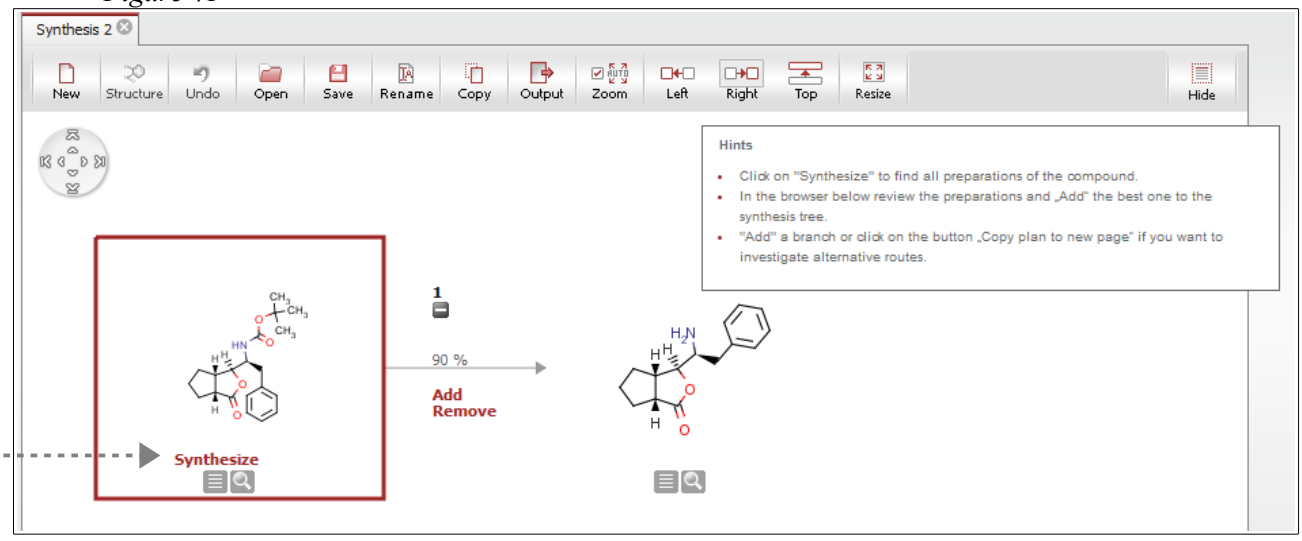
Ce tout nouvel outil apporté par Reaxys permet de consulter et modifier un plan de synthèse à partir des réactions décrites dans la base. On peut ouvrir **Synthesis Plans** directement à partir du bandeau supérieur sur une page vierge. Pour les premières utilisations, l'aide proposée à droite de la fenêtre (**Hints**) peut être intéressante.

L'autre manière d'utiliser les plans de synthèse se fait directement depuis une liste de réponses; supposons que l'on s'intéresse à la préparation du produit ci-dessous :



En activant le lien **Synthesize** sous le produit d'arrivée, la fenêtre des plans de synthèse s'ouvre et affiche la réaction demandée (Fig. 41) à partir de laquelle on va retrouver les réactions de préparation du produit de départ cette fois en activant à nouveau **Synthesize**.

Figure 41



Cette nouvelle action modifie la partie basse de la fenêtre, celle où apparaissent les réponses. Il y avait au départ une seule réaction; le fait de demander les réactions de préparation du précurseur conduit à l'affichage de 6 nouvelles réactions de préparation de ce composé (Fig. 41). Sur cette copie d'écran, seule la première de ces 6 réactions est montrée. On peut à ce niveau choisir quelle paraît la meilleure solution pour synthétiser cet intermédiaire; l'exemple est donné avec la première des 6 réactions.

Si on active **Add** (Fig. 42) on déclenche l'ouverture d'une nouvelle fenêtre (Fig. 43) affichant deux étapes supplémentaires car la réaction choisie était une multi-étapes. On aboutit ainsi à un arbre de synthèse qui se complexifie progressivement.

Figure 42

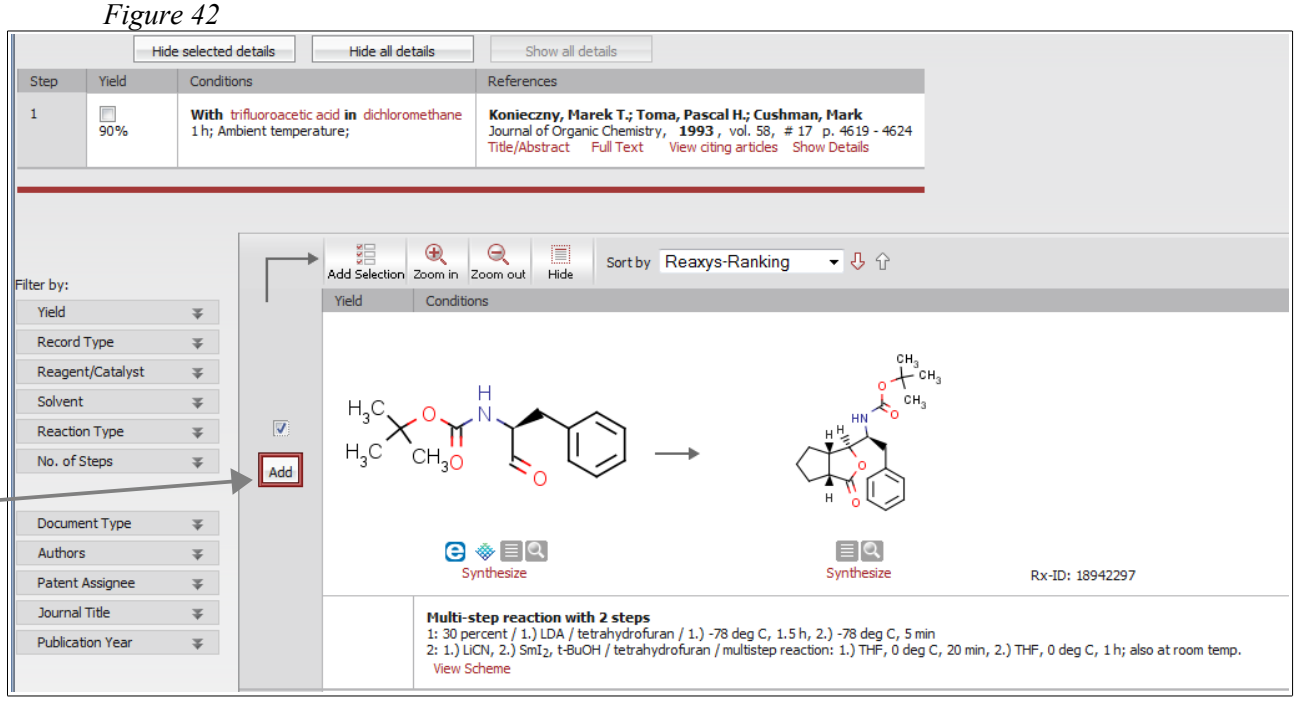
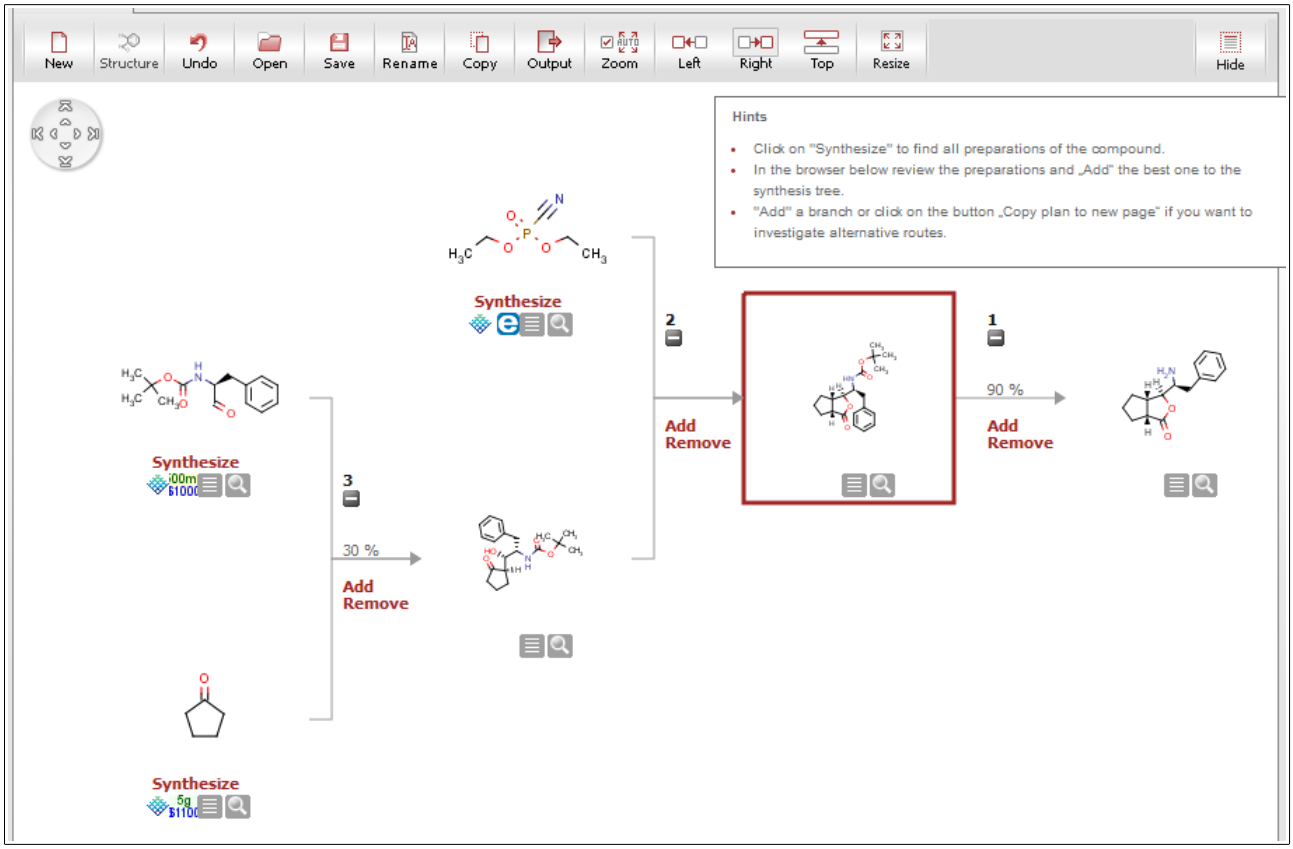


Figure 43



Remarque : on peut faire une réduction du plan avec le signe "-" pour réduire le plan lorsqu'il est très fourni.

Le lecteur est invité à découvrir les outils proposés par l'éditeur en haut de la fenêtre **Synthesis Plans** :

- l'outil **Copy** envoie une copie du plan en cours vers une autre page; cela peut être utile pour explorer

plusieurs voies de synthèse simultanément. Cette option de créer un nouveau plan de synthèse ne devient intéressante que si la réaction concernée a été décrite d'au moins deux manières différentes, voire par deux auteurs différents.

- l'outil **Output** exporte le plan en format pdf, vers un traitement de texte ou en format rdf (**reaction data file**) un format dans lequel les schémas réactionnels sont réexploitables graphiquement.
- les outils de zoom et en particulier **Resize** pour ajuster la taille de la fenêtre à celle de l'écran.

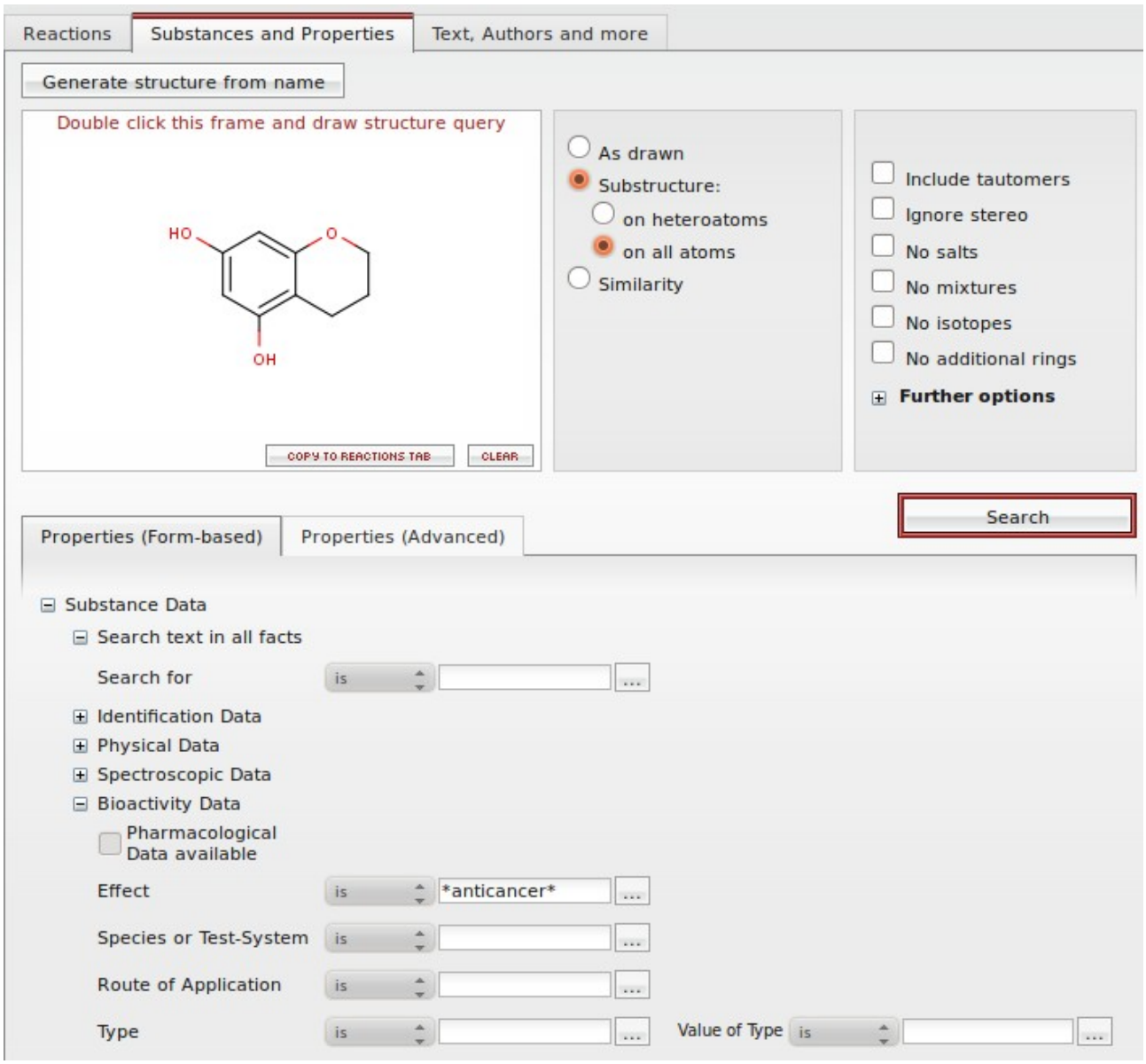
F – REQUÊTES CROISÉES

Nous avons vu dans les sections B, D et E comment interroger par Propriétés, Structures et Réactions respectivement. La possibilité d'effectuer des requêtes croisées est liée à l'interface de requête de Reaxys, car elle est divisée en 2 grandes parties :

- la partie structurale, que ce soit pour des réactions ou des substances, avec les options correspondantes
- la partie propriétés

On peut interroger simultanément ces deux parties; c'est d'ailleurs l'option par défaut. Dans l'exemple ci-dessous, la requête porte à la fois sur la sous-structure ET sur **ses** propriétés anticancéreuses. On a utilisé ici (Fig. 44) la partie **Properties (Form-based)** et le caractère générique "*" pour englober tous les termes contenant "anticancer" dans la requête.

Figure 44

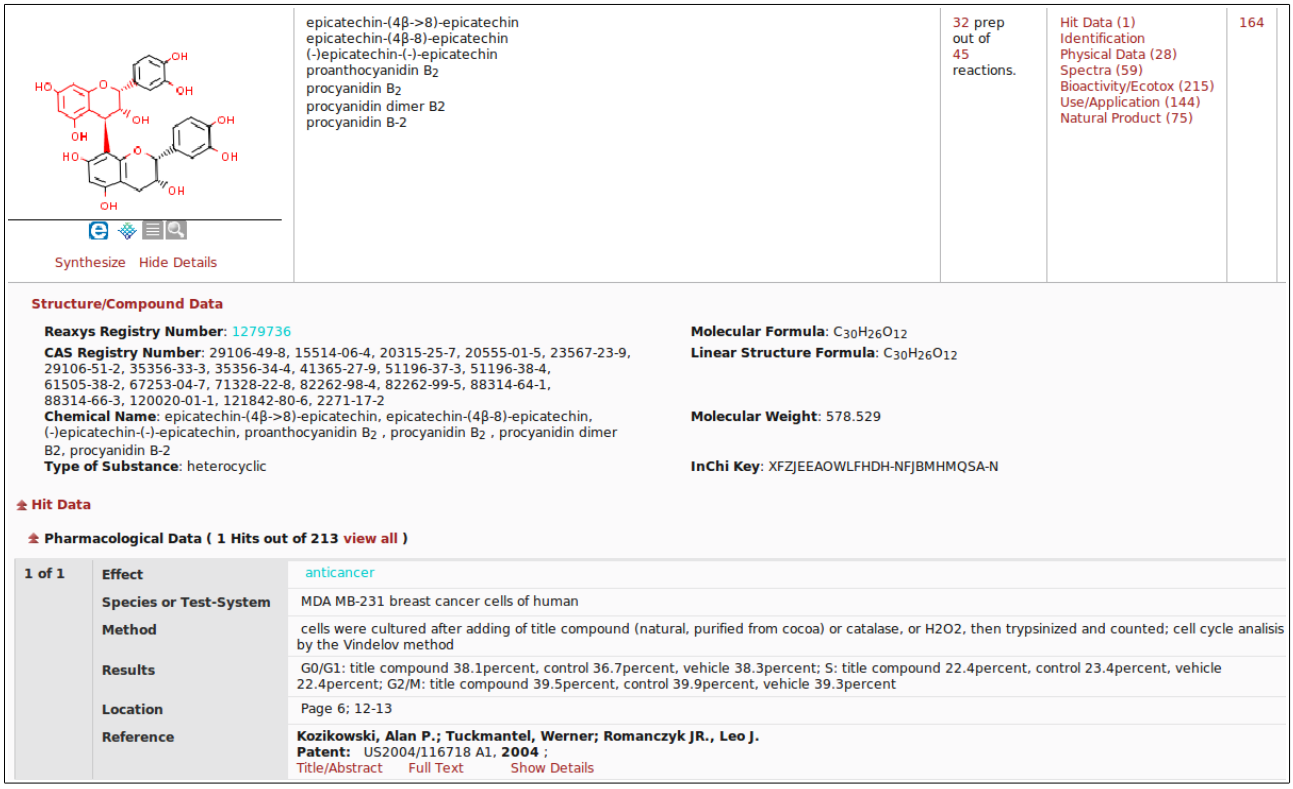


La question ci-dessus conduit à 44 réponses pertinentes dont une est montrée sur la Fig. 45. On notera la présence d'un troisième cycle, accolé ou non au bicycle recherché, l'option **No additional rings** n'étant pas activée.

On pourra imaginer des recherches plus précises encore en utilisant la recherche structurale ou sous-

structurale ET la partie **Properties (Advanced)**. Pour cette interrogation experte des propriétés, l'emploi des booléens **and**, **or, not, proximity, near** ou **next** pourra une grande précision dans les requêtes (cf. exemples en B-II-6)

Figure 45


epicatechin-(4β->8)-epicatechin
epicatechin-(4β-8)-epicatechin
(-)epicatechin-(-)-epicatechin
proanthocyanidin B2
procyanidin B2
procyanidin dimer B2
procyanidin B-2
32 prep out of 45 reactions.
Hit Data (1)
Identification
Physical Data (28)
Spectra (59)
Bioactivity/Ecotox (215)
Use/Application (144)
Natural Product (75)
164
Synthesize Hide Details
Structure/Compound Data
Reaxys Registry Number: 1279736
CAS Registry Number: 29106-49-8, 15514-06-4, 20315-25-7, 20555-01-5, 23567-23-9, 29106-51-2, 35356-33-3, 35356-34-4, 41365-27-9, 51196-37-3, 51196-38-4, 61505-38-2, 67253-04-7, 71328-22-8, 82262-98-4, 82262-99-5, 88314-64-1, 88314-66-3, 120020-01-1, 121842-80-6, 2271-17-2
Chemical Name: epicatechin-(4β->8)-epicatechin, epicatechin-(4β-8)-epicatechin, (-)epicatechin-(-)-epicatechin, proanthocyanidin B2 , procyanidin B2 , procyanidin dimer B2, procyanidin B-2
Type of Substance: heterocyclic
Molecular Formula: C30H26O12
Linear Structure Formula: C30H26O12
Molecular Weight: 578.529
InChi Key: XFZJEEAOWLFHDH-NFJBMHMQSA-N
Hit Data
Pharmacological Data (1 Hits out of 213 view all)
1 of 1
Effect
anticancer
Species or Test-System
MDA MB-231 breast cancer cells of human
Method
cells were cultured after adding of title compound (natural, purified from cocoa) or catalase, or H2O2, then trypsinized and counted; cell cycle analisis by the Vindelov method
Results
G0/G1: title compound 38.1percent, control 36.7percent, vehicle 38.3percent; S: title compound 22.4percent, control 23.4percent, vehicle 22.4percent; G2/M: title compound 39.5percent, control 39.9percent, vehicle 39.3percent
Location
Page 6; 12-13
Reference
Kozikowski, Alan P.; Tuckmantel, Werner; Romanczyk JR., Leo J.
Patent: US2004/116718 A1, 2004 ;
Title/Abstract Full Text Show Details

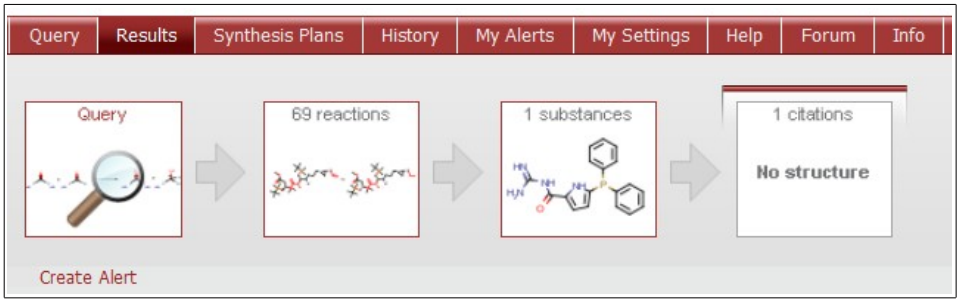
Remarque : on prendra garde à effacer tout ou partie de la requête si on veut démarrer une autre recherche.

G – TRAITEMENT DES RÉSULTATS

G-I- Navigation dans les résultats

Après avoir obtenu des résultats issus de contextes différents (réactions, substances, citations), on peut revenir en arrière et les consulter à nouveau grâce aux **breadcrumbs**[2]. illustrés sur la Fig. 46 ;

Figure 46



Cet exemple d'aide à la navigation provient des recherches successives à partir d'une liste de 69 réactions obtenue en E-IV à propos des réactions sélectives (2° carré), dont une substance – une guanidine substituée – a été sélectionnée (3° carré). Une des propriétés physiques de cette substance conduit à une citation particulière (4° carré). Chaque carré est évidemment un hyperlien.

G-II- Croisements de listes de réponses

Les croisements de liste permettent de réaliser des unions ou intersections ou exclusions de listes de réponses déjà obtenues. Ils permettent de comparer des listes issues du même contexte : réactions, substances, citations.

A titre d'exemple reprenons celui des réactions de type Friedel et Crafts vu au paragraphe E-II-2 ; selon que le mapping a été spécifié ou non sur le schéma réactionnel, nous avons obtenu 77 et 26 réponses respectivement. Nous pouvons visualiser ces deux ensembles de réponses *via* le menu **History** du bandeau principal (Fig. 47).

Pour connaître l'ensemble des réponses qui ont été écartées de l'ensemble des 26 réactions, on coche les cases correspondant aux listes de réactions (pas les citations!); l'outil **Combine hitsets** est alors actif et renvoie sur le multi-choix de la Fig. 48 où nous choisissons l'option **Exclude 2 from 4** pour n'obtenir que les 6 dibromo-pyridines recherchées.

G-III- Sauver/Retrouver une requête, une réponse

Les requêtes, les réponses et l'historique associé sont conservés par défaut sur le serveur jusqu'à déconnexion de la session. On peut toutefois les sauvegarder sur son poste de travail en passant par le mode **History** comme précédemment, auquel cas on pourra les récupérer ultérieurement.

Au cours d'une session, on peut consulter ses requêtes et listes de réponses mais leur nombre est limité à 20 requêtes par utilisateur stockées temporairement. L'intérêt de les sauvegarder sur son poste de travail est de pouvoir dépasser cette limite de 20 requêtes et de les consulter ultérieurement.

La sauvegarde des requêtes se fait par l'outil **Save Query** présent sur chaque interface de requête (Fig. 5). La sauvegarde des réponses se fait en mode **History** (Fig. 47) par l'outil **Store**.

G-IV- Veille scientifique

Cet outil s'adresse aux personnes qui veulent suivre un domaine de recherche au cours du temps. Après avoir posé une requête dans un des 3 contextes habituels, on passe en mode **History** pour générer une alerte avec l'outil **Create Alert** (Fig. 47).

2 - allusion aux miettes de pain laissées par le Petit Poucet dans la forêt pour retrouver son chemin ...

Figure 47

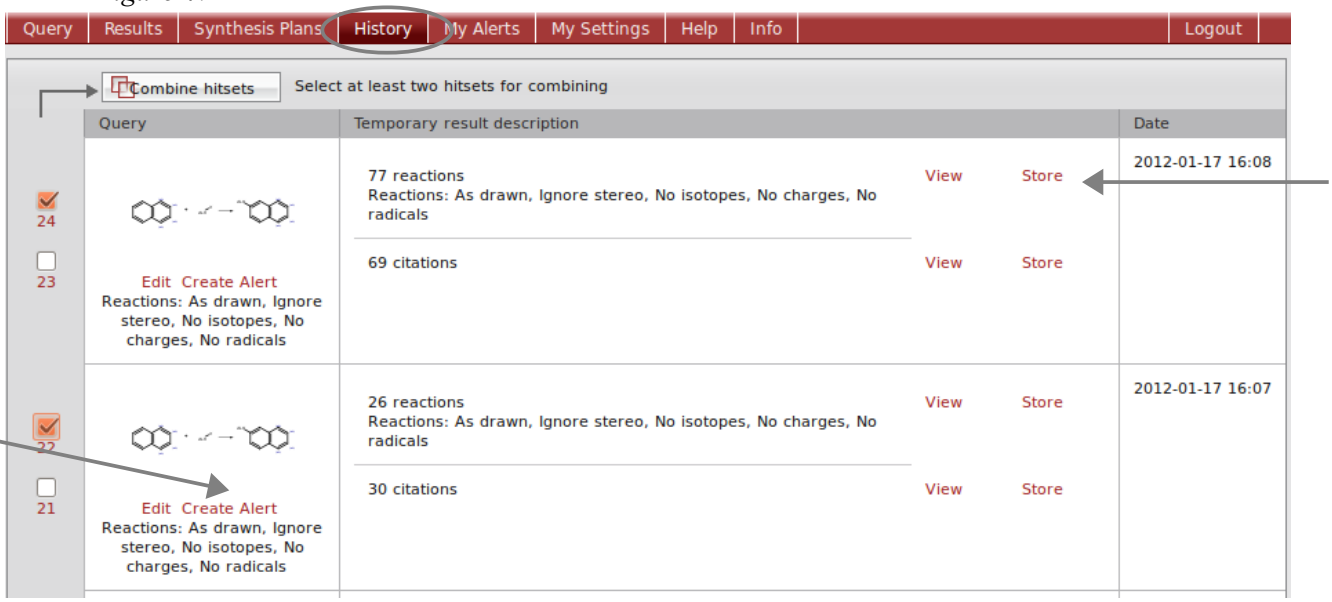


Figure 48

G-V- Export des résultats

Hormis les impressions classiques sur papier, de nombreuses possibilités sont offertes pour l'export des résultats. A partir d'une liste de réponses de substances, montrée partiellement sur la Fig. 49, la commande **Output** permet d'ouvrir la boîte de dialogue de la Fig. 50 où sont mentionnés les différents formats d'export.

Figure 49

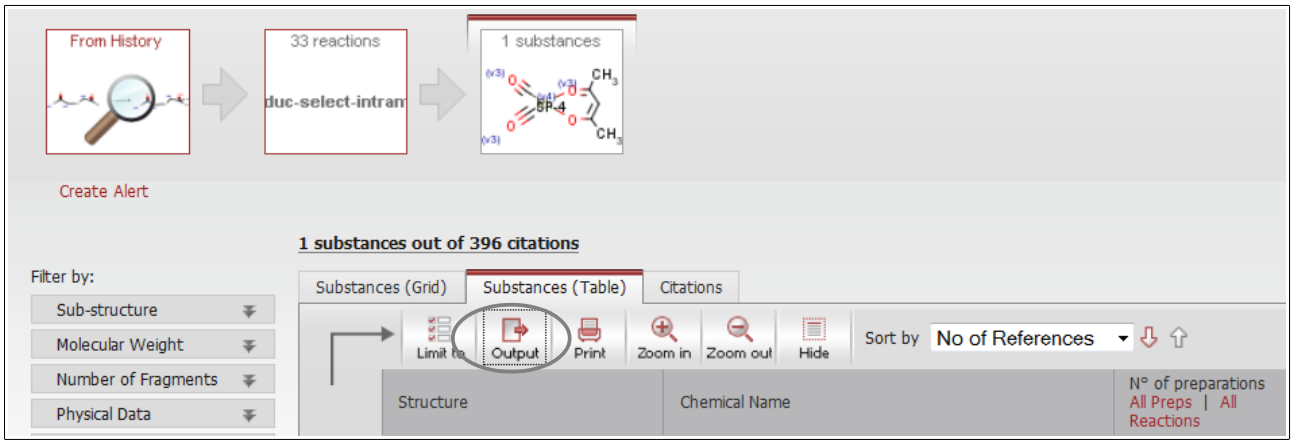
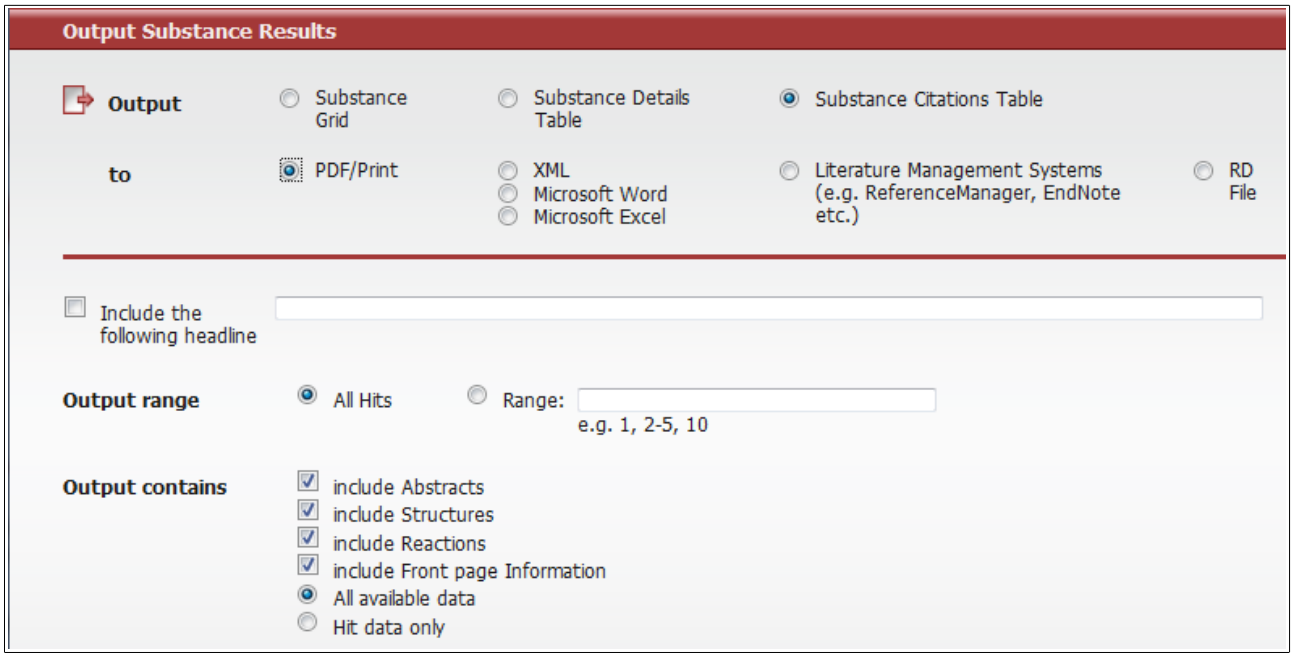


Figure 50



On notera que quel que soit le contexte des réponses (réactions, substances, citations), on peut aussi exporter les citations correspondantes dans un gestionnaire de références ; il suffira de cocher les cases **Substance Citations Table** puis **Literature Management Systems** dans l'exemple ci-dessus.

ANNEXES

Annexe 1 - Basic Indexes

1-1- Substance Basic Index (BISUB)

Le **Substance Basic Index** contient les champs suivants :

CN	Chemical Name
AUN	Autoname
CCN	Composition: Comp. Name
BRN	Beilstein Registry Number
RN	CAS Registry Number
INP	Isolation from Natural Product
RSTR	Related Structure
PUR	Purification
CNF.OBJ	Object of Investigation
EM.MET	Electrical Moment Method
IP.MET	Ionization Potential Method
CSYS	CSYS
NMR.NUC	NMR Nucleus
NMR.NUI	NMR Coupling Nuclei
ESR.NUI	ESR Coupling Nuclei
NQR.NUC	NQR Nucleus
DE.MET	Method
.SOL	All fields with the extension (except RX.SOL)
.KW / .COM	All fields with the extension
.CN / .PA	All fields with the extension

Type d'indexation : Chaîne de caractères

1-2- Reaction Basic Index (BIREA)

Le **Reaction Basic Index** contient les champs suivants :

RX.RCT	Reactant
RX.PRO	Product
RX.RGT	Reagent
RX.CAT	Catalyst
RX.SOL	Solvent
RX.COND	Other Conditions
RX.SUB	Subject Studied
RX.TYP	Reaction Type
RX.PRT	Prototype Reaction
RX.COM	Comment

Type d'indexation : Chaîne de caractères

1-3- Citation Basic Index (BICIT)

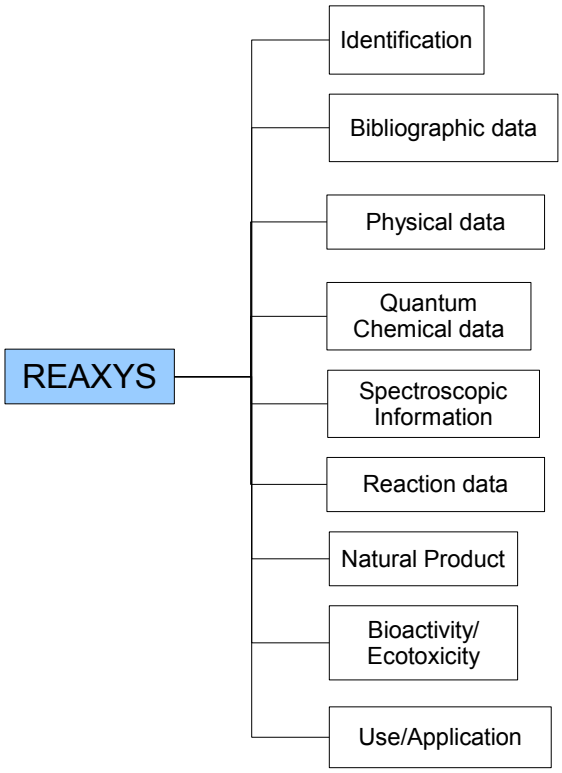
Le **Citation Basic Index** est seulement visible et disponible pour le sous-ensemble des **Abstracts databases**; il contient les champs suivants :

AU	Author
TI	Title
AB	Abstract
AB.KW	Abstract Keywords
JT	Journal Title
PA	Patent Assignee

Type d'indexation : Chaîne de caractères

Annexe 2 - Beilstein Data Structure

On trouvera ci-dessous un ensemble de tableaux récapitulant la structuration des données dans Reaxys. Chaque sous-ensemble de la base, montré sur le diagramme ci-dessous, est détaillé dans les pages suivantes.



Identification

Substance Identification

Reaxys Registry Number (IDE.XPR)
Preferred Registry Number (IDE.XRN)
CAS Registry Number (IDE.RN)
Chemical Name (IDE.CN)
Chemical Name Segment (IDE.CNS)
Linear Structure Formula (IDE.LSF)
Molecular Formula (IDE.MF)
Search MF Range (IDE.MOFO)
Charge (IDE.CHA)
Element Counts (IDE.ELC)
Number of Atoms (IDE.NA)
Number of Elements (IDE.NE)
Number of Fragments (IDE.NF)
Molecular Weight (IDE.MW)
Inchi Key (IDE.INCHI)

One-Center Ligands

Base Formula (LIGO.BASE)
Formula (LIGO.FORM)
Count (LIGO.CNT)

Multi-Center Ligands

Base Formula (LIGM.BASE)
Formula (LIGM.FORM)
Count (LIGM.CNT)

Alloy Composition

Formula (ALLOYFORM)

Patent Specific Data

Prophetic Compound (PSD.PRC)
Related Markush Structure (PSD.MARPRN)
Location in Patent (PSD.LCN)

Related Structure (RSTR.RSTR)

Derivative(CDER.CDER)

Derivative Meting Point, °C (CDER.MP)

Purification (PUR.PUR)

Natural Product

Isolation from Natural Product (INP.INP)

Spectroscopic Information

NMR Spectroscopy

Description (NMR.KW)
Nucleus (NMR.NUC)
Coupling Nuclei (NMR.NUI)
Solvent (NMR.SOL)
Temperature, °C (NMR.T)
Frequency (NMR.F)
Original text (NMR.TXT)

IR Spectroscopy

Description (IR.KW)
Solvent (IR.SOL)
Temperature, °C (IR.T)
Temperature, °C (IR.T)

Mass Spectrometry

Keywords (MS.KW)

UV/VIS Spectroscopy

Description (UV.KW)
Solvent (UV.SOL)
Absorption maxima (UV.AM)
Ext./Abs. Coefficient, $l.mol^{-1}.cm^{-1}$ (UV.EAC)

ESR Spectroscopy

Description (ESR.KW)
Coupling nuclei (ESR.NUI)
Solvent (ESR.SOL)
Temperature, °C (ESR.T)

NQR Spectroscopy

Description (NQR.KW)
Nucleus (NQR.NUC)

Rotational Spectroscopy

Description (ROT.KW)

Raman Spectroscopy

Description (RAMAN.KW)
Solvent (RAMAN.SOL)

Luminescence Spectroscopy

Description (LUM.KW)

Fluorescence Spectroscopy

Description (FLU.KW)
Solvent (FLU.SOL)
Temperature, °C (FLU.T)

Phosphorescence Spectroscopy

Description (PHO.KW)
Solvent (PHO.SOL)
Temperature, °C (PHO.T)

Other Spectroscopic Methods

Description (OSM.KW)
Nucleus (OSM.MOENUC)

Bibliographic Data

Citation

Citation Number (CNR.CNR)
Citation Document Type (CIT.DT)
Author (CIT.AU)
Patent assignee (CIT.PA)
Common Patent Number (CIT.PNX)
Patent Country code (CIT.PCC)
Patent Number (CIT.PN)
Patent Year (CIT.PPY)
Patent Status (CIT.PK)
CODEN (CIT.CO)
Journal Title (CIT.JT)
Journal Title (Short) (CIT.JTS)
Journal/Review (CIT.JTW)
Country code (CIT.CC)
Language Code (CIT.LA)
Publisher (CIT.PUB)
Location of Publisher (CIT.LO)
(Series) Volume (CIT.VL)
Number (CIT.NB)
Publication Year (CIT.PY)
Page (CIT.PAG)
Citation Unresolved (CIT.URES)
DOI (CIT.DOI)
Article Number (CIT.ARTNO)
ISSN (CIT.ISSN)

Abstract

Title (AB.TI)
Abstract (AB.AB)
Language (AB.ALA)
Keywords (AB.KW)

Patent bibliography

Date of publication (PBIB.PD)
Application number (PBIB.AP)
Date of Filing (PBIB.FD)
Manually excerpted (PBIB.MAN)
Inventor (PBIB.INV)
Priority number (PBIB.PRNR)
Priority date (PBIB.PRD)
Main IPC (PBIB.ICM)
Secondary IPC (PBIB.ICS)
Family member: Patent number (PBIB.FPN)
Family member: Status code (PBIB.FPSTAT)
Family member: Date of publication (PBIB.FPD)
Family member: Application number (PBIB.FAP)
Family member: Date of filing (PBIB.FFD)
Family member: Indexed patent (PBIB.FIDX)

Reaction Data

Reaction

Reaction ID (RX.ID)
Reactant XRN (RX.RXRN)
Reactant (RX.RCT))
Product XRN (RX.PXRN))
Product (RX.PRO))
Number of Reaction Details (RX.NVAR))
Record type (RX.RTYP)

Reaction Details

Reaction Classification (RXD.CL)
Full text of Reaction (RXD.TXT)
Number of Reaction Steps (RXD.STP)
Product XRN (RXD.YXRN)
Product (RXD.YPRO)
Yield (RXD.YD)
Yield (Numerical) (RXD.NYD)
Yield (optical) (RX.YDO)
Number of Stages (RXD.SNR)
Solvent (RXD.SOL)
Reagent/Catalyst (RXD.RGTCAT)
Time (RXD.TIM)
Temperature, °C (RXD.T)
Pressure, Torr (RXD.P)
pH value (RXD.PH)
Other Conditions (RXD.COND.)
Reaction Type (RXD.TYP)
Subject Studied (RXD.SUB)
Prototype Reaction (RXD.PRT)

Reaction Basic Index

Reaction Basic Index (RBI.BIREA)

Physical Data (PHY)

Melting Point (MP.MP)
- Solvent (MP.SOL)

Boiling Point (BP.BP)
- Pressure, Torr (BP.P)

Sublimation (SP.SP)
- Pressure, Torr (SP.P)

Refractive Index (RI.RI)
- Wavelength (RI.W)
- Temperature, °C (RI.T)

Density of the Liquid (DEN.DEN)
- Reference Temperature, °C (DEN.RT)
- Measurement Temperature (DEN.T)

Adsorption (MCS) (ADSM.KW)
- Partner XRN (ADSM.PB)
- Partner (ADSM.PA)
- Solvent (ADSM.SOL)
- Temperature, °C (ADSM.T)
- Pressure, Torr (ADSM.P)

Association (MCS) (ASSM.KW)
- Partner XRN (ASSM.PB)
- Partner (ASSM.PA)
- Solvent (ASSM.SOL)
- Temperature, °C (ASSM.T)
- Pressure, Torr (ASSM.P)

Autoignition Temperature (AUTI.T)

Azeotropes (MCS)
- Azeotropes (AZE.PA)
- Temperature, °C (AZE.T)
- Pressure, Torr (AZE.P)
- Concentrations (AZE.C)

Boundary Surface Phenomena (MCS) (BSPM.KW)
- Partner (BSPM.PA)
- Solvent (BSPM.SOL)
- Temperature, °C (BSPM.T)
- Pressure, Torr (BSPM.P)

Bulk Viscosity (BV.BV)
- Temperature, °C (BV.T)

Circular Dichroism
- Solvent (CDIC.SOL)

Complex Phase Equilibria (MCS) (CPEM.KW)
- Partner XRN (CPEM.PB)
- Partner (CPEM.PA)
- Solvent (CPEM.SOL)
- Temperature, °C (CPEM.T)

Compressibility (COMP.KW)

Conformation
- Object of investigation (CNF.OBJ)

Critical Density (CRD.CRD)

Critical Micelle Concentration (MCS)
- Critical Micelle Concentration (CMC.CMC)
- Temperature, °C (CMC.T)
- Solvent (CMC.SOL)

Critical Pressure (CRP.CRP)

Critical Temperature (CRT.CRT)

Critical Volume (CRV.CRV)

Crystal Phase (CRYPH.KW)
- Temperature, °C (CRYPH.T)

Crystal Property Description Colour & Other Properties (CPD.CPD)
- Point Group (CPD.PGROUP)

Crystal System (CSYS.CSYS)

Decomposition (DP.DP)
- Solvent (DP.SOL)
- Solvent for Crystallization (DP.CRSOL)
- Amount (DP.SOLM)

Density of the Crystal (CDEN.CDEN)
- Temperature, °C (CDEN.T)

Dielectric Constant (DIC.DIC)
- Frequency (DIC.F)
- Temperature, °C (DIC.T)

Dissociation Energy (EDIS.EDIS)
- Bond Type (EDIS.TYP)

Dissociation Exponent (pK) (DE.DE)
- Dissociation Group (DE.GRP)
- Temperature, °C (DE.T)
- Solvent (DE.SOL)
- Method (DE.MET)
- Type (DE.TYP)

Dynamic Viscosity (DV.DV)
- Temperature, °C (DV.T)

Electrical Data (ELE.KW)
- Electric Conductivity (ELE.ECVAL)
- Temperature of Electric Conductivity, °C (ELE.T)
- Critical Superconductivity Temperature, °C (ELE.CRIT)

Electrical Data (MCS) (EDM.KW)
- Partner XRN (EDM.PB)
- Partner (EDM.PA)
- Solvent (EDM.SOL)
- Temperature, °C (EDM.T)
- Pressure, Torr (EDM.P)

Electric Moment
- Electric Moment (EM.KW)
- Moment (EM.EM)
- Temperature, °C (EM.T)
- Method (EM.MET)
- Solvent (EM.SOL)

Electrical Polarizability (ELP.KW)

Electrochemical Behaviour (ELCB.KW)

Electrochemical Characteristics (POT.KW)
- Solvent (POT.SOL)
- pH-Value (POT.PH)
- Temperature, °C (POT.T)

Electrochemistry Data (ELCH.KW)
- Electrochemical Cell (ELCH.DESCR)
- Cell Potential, V (ELCH.POT)

Electrolytic Conductivity (ELYC.KW)
- Electrolytic Conductivity (ELYC.VAL)
- Equivalent Conductivity (ELYC.AEVAL)
- Temperature, °C (ELYC.T)
- Solvent (ELYC.SOL)
- Kind of Conductivity (ELYC.REM)

Electron Binding (CIP.KW)

Energy Barriers (EBC.EBC)
- Barrier Type (EBC.TYP)
- Solvent (EBC.SOL)

Energy Data (MCS) (ENEM.KW)
- Partner XRN (ENEM.PB)
- Partner (ENEM.PA)
- Solvent (ENEM.SOL)
- Temperature, °C (ENEM.T)
- Pressure, Torr (ENEM.P)

Enthalpies of Other Phase Transitions (HPTP.HPTP)

Enthalpy of Combustion (HCOM.HCOM)
- Temperature, °C (HCOM.T)
- Pressure, Torr (HCOM.P)

Enthalpy of Formation (HFOR.HFOR)
- Temperature, °C (HFOR.T)
- Pressure, Torr (HFOR.P)

Enthalpy of Fusion (HFUS.HFUS)

Enthalpy of Hydrogenation (HHDG.HHDG)
- Product XRN (HHDG.XRN)
- Product Name (HHDG.CN)
- Temperature, °C (HHDG.T)

Enthalpy of Sublimation (HSP.HSP)
- Temperature, °C (HSP.T)

Enthalpy of Vaporization (HVAP.HVAP)
- Temperature, °C (HVAP.T)
- Pressure, Torr (HVAP.P)

Explosion Limits (EXPL.LV)

Flash Point
- Temperature, °C (FLAP.FLAP)
- Type of Test (FLAP.TYP)

Further Information (FINFO.KW)
- Additional Description (FINFO.KW2)

Gas Phase (GP.KW)

Heat Capacity Cp (CP.CP)
- Temperature, °C (CP.T)

Heat Capacity Cp0 (CP0.CP0)
- Temperature, °C (CP0.T)

Heat Capacity Cv (CV.CV)
- Temperature, °C (CV.T)

Henry Constant (MCS)
- Henry Constant (HEN)
- log Henry Constant (HEN.LOG)
- Temperature, °C (HEN.T)
- Solvent (HEN.SOL)

Interatomic Distances and Angles (IDA.KW)

Ionization Potential (IP.IP)
- Method (IP.MET)

Isoelectric Point pH (IEP)
- Solvent (IEP.SOL)

Kinematic Viscosity (KV.KV)
- Temperature, °C (KV.T)

Liquid Phase (LIQPH.KW)

Liquid/Liquid Systems (MCS) (LLSM.KW)
- Partner (LLSM.PA)
- Solvent (LLSM.SOL)
- Temperature, °C (LLSM.T)
- Pressure Description (LLSM.P)

Liquid/Solid Systems (MCS) (LSSM.KW)
- Partner (LSSM.PA)
- Solvent (LLSM.SOL)
- System Component (LLSM.SOL)
- Temperature, °C (LSSM.T)
- Pressure, Torr (LLSM.P)

Liquid/Vapour Systems (MCS) (LVSM.KW)
- Partner (LVSM.PA)
- Solvent (LVSM.SOL)
- Temperature, °C (LVSM.T)
- Pressure, Torr (LVSM.P)

Magnetic Data (MAG.KW)
- Temperature, °C (MAG.T)
- Moment (MAG.MMOM)

Physical Data (PHY)

Magnetic Susceptibility (MSUS.MSUS)
- Temperature, °C (MSUS.T)

Mechanical & Physical Properties (MCS) (MECM.KW)
- Partner (MECM.PA)
- Solvent (MECM.SOL)
- Temperature, °C (MECM.T)
- Pressure, Torr (MECM.P)

Mechanical Properties (MEC.KW)

Molecular Deformation (DFM.KW)

Mutarotation Type (MUT.TYP)
- Concentration (MUT.C)
- Length of Path (MUT.LEN)
- Solvent (MUT.SOL)
- Mutarotation (MUT.MUT)
- Wavelength (MUT.W)
- Temperature, °C (MUT.T)
- Time (MUT.TIM)

Optical Data (MCS) (ODM.KW)
- Partner XRN (ODM.PB)
- Partner (ODM.PA)
- Solvent Temperature, °C (ODM.T)
- Pressure, Torr (ODM.P)

Optical Rotatory Dispersion Solvent (ORD.SOL)

Optical Rotatory Power (ORP.TYP)
- Concentration (ORP.C)
- Length of Path (ORP.LEN)
- Solvent (ORP.SOL)
- Optical Rotatory Power (ORP.ORP)
- Wavelength (ORP.W)
- Temperature, °C (ORP.T)

Optics (OPT.KW)

Other Thermochemical Data (OTHE.KW)

Pressure (CPEM.P)

Partition octan-1-ol/Water (MCS)
- Partition Constant (POW)
- log POW (POW.LOG)
- Temperature, °C (POW.T)

Self-diffusion
- Self-diffusion Coefficient (SDIF.SDIF)
- Temperature, °C (SDIF.T)

Solubility (MCS)
- Solubility (SLB)
- Saturation (SLB.SAT)
- Temperature, °C (SLB.T)
- Solvent (SLB.SOL)
- Ratio of Solvents (SLB.RAT)

Solubility Product (MCS)
- Solubility Product 'SLBP)
- Temperature, °C (SLBP.T)
- Solvent (SLBP.SOL)
- Ratio of Solvents (SLBP.RAT)

Solution Behaviour (MCS) (SOLM.KW)
- Solvent (SOLM.SOL)
- Temperature, °C (SOLM.T)
- Pressure, Torr (SOLM.P)

Sound Properties (SOUND.KW)

Space Group Description (CSG.CSG)

Static Dielectric Constant (SDIC.SDIC)
- Temperature, °C (SDIC.T)

Surface Tension (ST.ST)
- Temperature, °C (ST.T)

Thermal Expansion
- Coefficient (TEXP.VAL)
- Kind of Expansion (TEXP.LEC)
- Temperature, °C (LEC.T)

Transition Point(s) of Crystalline Modification(s)
- Temperature, °C (CPTP.CPTP)
- Change of Modification (CPTP.CM)

Transition Point(s) of Liquid Modification(s) Temperature (LPTP.LPTP)
- Change of Modification (LPTP.CM)

Transport Data (TD.KW)

Transport Phenomena (MCS) (TRAM.KW)
- Partner (TRAM.PA)
- Solvent (TRAM.SOL)
- Temperature, °C (TRAM.T)
- Pressure, Torr (TRAM.P)

Triple Point (TP.TP)

Vapour Pressure (VP.VP)
- Temperature, °C (VP.T)

Quantum Chemical Data

- Calculated Properties (QUAN.PROP)
- Method (QUAN.MET)

Bioactivity/Ecotoxicology

Pharmacological Data

Effect (PHARM.E)
Endpoint of Effect (PHARM.EP)
Species or Test-System (PHARM.SP)
Sex (PHARM.S)
Route of Application (PHARM.RA)
Concentration (PHARM.C)
Kind of Dosing (PHARM.KD)
Exposure Period (PHARM.EX)
Method (PHARM.MR)
Further Details (PHARM.FD)
Half-Life Time (PHARM.H)
Type (PHARM.TY)
Value of Type (PHARM.V)
Results (PHARM.RE)
Metabolite (PHARM.META)
Comment (PHARM.COM)

Ecological Data

Ecotoxicology

Effect (ECT.E)
Endpoint (ECT.EP)
Species or Test-System (ECT.SP)
Sex (ECT.S)
Route of Application (ECT.RA)
Concentration (ECT.C)
Kind of Dosing (ECT.KD)
Exposure Period (ECT.EX)
Method (ECT.MR)
Further Details (ECT.FD)
Type (ECT.TY)
Value of Type (ECT.,V)
Results (ECT.RE)
Metabolite (ECT.META)

Exposure Assessment

Exposure (ECA.HE)
Sources (ECA.SO)

Concentration in the Environment

Species (ECC.SP)
Media (ECC.ME)
Location (ECC.LO)
Contamination Concentration (ECC.CC)
Background Concentration (ECC.BC)
Method (ECC.MR)

Transport and Distribution (ECTD)

Type (ECTD.TY)
Media (ECTD.ME)
Results (ECTD.RE)
Method (ECTD.MR)

Bioaccumulation, Biomagnification and Biomonitoring

Species (BIO.SP)
Medium (BIO.ME)
Concentration (BIO.C)
Exposure Period (BIO.E)
Temperature, °C (BIO.T)
Bioconcentration Factor (BIO.BC)
Log BCF (BIO.LOG)
Accumulation Half-life Time (BIO.A)
Accumulation Rate Constant (BIO.AR)
Elimination Half-Life Time (BIO.H)
Elimination Rate Constant (BIO.ER)
Method (BIO.MR)
Biomagnification (BIO.MAG)
Biomonitoring (BIO.MON)

Biodegradation (BIOD)

Type (BIOD.TY)
Inoculum (BIOD.IN)
Concentration (BIOD.C)
Degradation Rate (BIOD.D)
Exposure Period (BIOD.E)
Temperature, °C (BIOD.T)
Half-Life Time (BIOD.H)
Degradation Product BRN (BIOD.BRN)
Degradation Product (BIOD.DP)
Method (BIOD.MR)

Abiotic Degradation, Hydrolysis (ECDH)

Type (ECDH.TY)
Concentration (ECDH.C)
Degradation Rate (ECDH.D)
Exposure Period (ECDH.E)
Temperature, °C (ECDH.T)
Rate Constant (ECDH.RC)
pH-Value (ECDH.PH)
Half-Life Time (ECDH.H)
Degradation Product (ECDH.DP)
Method, Remarks (ECDH.MR)

Abiotic Degradation, Photolysis (ECDP)

Type (ECDP.TY)
Concentration (ECDP.C)
Degradation Rate (ECDP.D)
Exposure Period (ECDP.E)
Temperature, °C (ECDP.T)
Rate Constant (ECDP.RC)
pH-Value (ECDP.PH)
Half-Life Time (ECDP.H)
Degradation Product (ECDP.DP)
Method (ECDP.MR)

Stability in Soil (ECS)

Type (ECS.TY)
Concentration (ECS.C)
Dissipation Time 50 (ECS.5)
Dissipation Time 90 (ECS.9)
Dissipation (ECS.D)
Exposure Period (ECS.E)
Temperature, °C (ECS.T)
pH-Value (ECS.PH)
Humidity (ECS.HU)
Organic Carbon (ECS.OC)
Cation Exchange Rate (ECS.CE)
Microbial Biomass (ECS.MB)
Method (ECS.MR)

Oxygen Demand (EOD)

Type (EOD.TY)
Related to (EOD.RE)
Concentration (EOD.C)
Oxygen Demand (EOD.D)
Ratio BOD5/COD (EOD.RAT)
Method (EOD.MR)

Use/Application

Laboratory Use and Handling (USE.LH)
Use Pattern (USE.PT)

INDEX

A

and.....
requêtes croisées 53
utilisation 19
atome.....
dessiner 23
Attachments.....
deux points d'attachement 33
un point d'attachement 31
AutoSearch ON..... 39

B

Basic Indexes..... 18
bicit..... 18
définition 58
birea..... 18
définition 57
recherche par faits 17
rétrosynthèse 48
structuration de la base 58
bisub..... 18
définition 57
booléens..... 18
brevets..... 3, 7

C

catalyseur.....
indexation de 17
centre réactionnel..... 44
chaîne.....
de taille variable 29
citantes.....
références 7
citations.....
export 56
connexion.....
au serveur 8
contexte.....
des Basic Indexes 18
des requêtes 18
cycle.....
de taille variable 29
nombre de 39

D

déconnexion.....
du serveur 8
dessiner.....
un atome 23
un schéma réactionnel 42
une liaison 25
domaine, de recherche..... 10

E

export vers.....
Excel 55

F

formulaires prédéfinis..... 12
Full Text..... 7

G

groupements.....
définis par l'utilisateur 31
génériques 31
variables 31
variables, structure de Markush 31
groupements fonctionnels..... 27

H

half reaction..... 5, 42
historique des requêtes..... 37, 55

I

indexation.....
problèmes 7, 8, 17
interrogation.....
par propriétés 12
inversion .....
de la stéréochimie 45

L

liaison.....
cassée/formée 44
dessiner 25
Link Node..... 29
liste d'atomes.....
définir 24
listes de réponses.....
croisements 54

M

mapping..... 44
Markush..... 4, 39
fréquence de 31, 34
mots-clés.....
recherche par 16

N

nombre.....
de cycles 39
nombre .....
de substituants 24, 38
not.....
requêtes croisées 53
utilisation 19

O

opérateurs logiques..... 18
or

requêtes croisées *53*
utilisation *19*

P

polymère..... 5
propriétés.....
contenu de la base *3*
écologiques *7*
pharmacologiques *7*
structuration de la base *4*
toxicologiques *7*

R

réaction.....
de Friedel et Crafts *43*
dessiner *42*
half reaction *42*
recherche par *42*
réactions.....
multi-étapes *47*
sélectives *46*
Reaxys Generics..... 35
recherche automatique..... 39
recherche, domaine de..... 10
références.....
citantes *7*
exporter une liste de *56*
rétention.....
de la stéréochimie *45*

S

sélectivité.....
au cours d'une réaction *46*
sous-structure.....
recherche par *38*
stéréochimie.....
liaison *27*
structure exacte.....
recherche par *37*
structures de Markush..... 31
substitution.....
précision sur les atomes *24, 38*

T

Templates..... 35
topologie chaîne/cycle..... 28
troncature..... 16